Centre de recherche sur les transports (C.R.T.)
Centre for Research on Transportation

CRT-2005-14

RECHERCHE OPÉRATIONNELLE ET TRANSPORT DE MARCHANDISES

par

Teodor Gabriel Crainic
Frédéric Semet

Mai 2005

Université de Montréal
C.P. 6128, succursale Centre-ville, Montréal QC H3C 3J7 Canada
Téléphone : (514) 343-7575 / Télécopieur : (514) 343-7121

crt@crt.umontreal.ca
www.crt.umontreal.ca

HEC ÉCOLE POLYTECHNIQUE MONTRÉAL Université de Montréal

105-05346/

PUBLICATION
CRT-2005-14

RECHERCHE OPÉRATIONNELLE ET TRANSPORT DE MARCHANDISES

par

Teodor Gabriel Crainic [1]
Frédéric Semet [2]

Ce projet a été financé en partie par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG).

[1] Département de management et technologie, École des sciences de la gestion, Université du Québec à Montréal, C.P. 8888, succursale Centre-ville, Montréal, Canada H3C 3P8 et Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Canada H3C 3J7 (theo@crt.umontreal.ca)

[2] LAMIH, Université de Valenciennes et du Hainaut-Cambrésis, Le Mont-Houy, 59311 Valenciennes Cedex 9, France (Frederic.Semet@univ-valenciennes.fr)

Dépôt légal – Bibliothèque nationale du Québec
Bibliothèque nationale du Canada, 2005

Abstract

The goal of this paper is to illustrate the richness of the applications of Operations Research to freight transportation problems. The focus is on a number of major problem classes with significant impact both in terms of applications and methodological development : network design, location, service network design, fleet management, vehicle routing problems. We present the issues, describe the models proposed, and discuss algorithmic and computational issues.

Keywords : freight transportation, models and methods, network design, location, service network design, fleet management, vehicle routing problems.

Résumé

L'objectif de ce rapport est d'illustrer la richesse des applications « transport des marchandises » de la recherche opérationnelle en ciblant quelques secteurs majeurs, à la fois par leur importance socio-économique et par les développements méthodologiques auxquels ils sont associés : conception des systèmes de transport et logistiques, conception des réseaux de service et gestion de flottes des transporteurs longue distance, ainsi que les problèmes de tournées de véhicules en se concentrant sur trois variantes principales : contraintes de capacité, contraintes de capacité et de longueur, et fenêtres de temps. Nous présentons les problèmes, décrivons les modèles associés et discutons de leur résolution.

Mots-clés : transport de marchandises, modèles et méthodes, conception de réseaux, localisation, conception de réseaux de service, gestion de flottes, problèmes de tournées de véhicules

1 Introduction

Le transport, et le transport des marchandises en particulier, constitue une composante vitale de la société humaine. Qu'il s'agisse d'activités économiques, sociales, culturelles ou autres, on retrouve la dimension « transport » parmi les éléments essentiels à leur réalisation. Au niveau économique, les activités de transport sont ainsi responsables d'une part importante du produit national brut d'un pays (de l'ordre de 10%), tandis que le coût de transport contribue significativement au prix de vente final des divers produits et services (de l'ordre de 10% à 15%). Par ailleurs, le transport est aussi responsable d'une proportion significative des émissions de gaz à effet de serre et, plus généralement, de la pollution de l'environnement,sans oublier les niveaux de congestion chronique de l'ensemble des infrastructures de transport auxquels font face la plupart des nations.

Les tendances économiques et sociales observées mondialement pointent vers l'augmentation continue et forte du niveau des transports. Dans le cas du transport des marchandises, cette tendance se manifeste à tous les niveaux, depuis le commerce international jusqu'à la livraison locale dans un quartier. Ce phénomène s'accompagne d'exigences accrues pour que les activités de transport soient rapides, fiables, sures et peu onéreuses. L'ensemble de ces facteurs met fortement en évidence l'importance d'une planification et d'une gestion adéquate des activités et systèmes de transport.

Il y a une longue tradition d'échanges fructueux entre la recherche opérationnelle et le transport. Les problèmes d'analyse, de planification et de gestion des systèmes de transport ont servi de point de départ et de défis pour un grand nombre de développements méthodologiques, du problème et algorithme « de transport » aux méta-heuristiques pour l'optimisation combinatoire. D'un autre côté, la recherche opérationnelle a permis d'apporter des solutions à un grand nombre de problèmes de planification et de gestion des systèmes de transport. Les modèles et méthodes de recherche opérationnelle se retrouvent au cœur de logiciels d'aide à la décision pour la planification des réseaux de transport urbains, régionaux ou internationaux, la conception des réseaux de service de transport pour les passagers comme pour les marchandises, la gestion des flottes de véhicules, la planification et la gestion des activités logistiques et de distribution, la construction des horaires et des routes des équipages et des véhicules, et ainsi de suite. Cette tradition se poursuit avec vigueur ainsi qu'en témoigne ce chapitre.

Le transport est un domaine vaste et varié auquel participe un grand nombre d'acteurs, chacun avec des objectifs propres et des problèmes d'analyse, de planification et de gestion spécifiques. Notre intention n'est pas de faire une présentation détaillée des activités et systèmes de transport. Une telle démarche déborderait largement du cadre de ce chapitre. Notre objectif est plutôt d'illustrer la richesse des applications « transport » de la recherche opérationnelle.

Nous ciblons le transport des marchandises, dont la section 2 présente une brève introduction, et traitons quelques secteurs majeurs à la fois par leur importance socio-

économique et par les développements méthodologiques auxquels ils sont associés. Nous avons structuré la présentation selon quelques grandes thématiques qui reflètent à la fois les niveaux de planification des opérations de transport et les étapes des chaînes logistiques de distribution. La section 3 traite de quelques modèles importants de conception des systèmes de transport et logistiques, qu'ils soient déployés à grande (région, nation, monde) ou petite (quartier, ville, région) échelle. La section 4 est dédiée aux problèmes de planification tactique et opérationnelle des transporteurs longue distance. On y traite, en particulier, des modèles de conception des réseaux de services, ainsi que de modèles visant la gestion de flottes. La section 5 traite un des problèmes centraux de transport, de distribution et de logistique, aussi relié à la gestion opérationnelle des flottes de véhicules, à savoir le problème de tournées de véhicules. Nous focalisons trois variantes principales de ce problème en considérant des contraintes de capacité, de capacité et de longeur, ainsi que de fenêtres de temps. La section 6 décrit les principaux modèles pour ces variantess ainsi que les méthodes de résolution exacte associées. La section 7 est quant à elle consacrée aux approches heuristiques. Nous concluons à la section 8. L'annexe contient une brève présentation de deux classes de méthodes méta-heuristique, la recherche tabou et les algorithmes évolutionnaires, qui sont largement utilisés afin d'identifier, avec un effort de calcul raisonnable, de bonnes solutions approchées aux problèmes combinatoires difficiles rencontrés dans presque tous les aspects de la planification et de la gestion des systèmes et activités de transport de marchandises.

2 Systèmes de transport de marchandises

Dans le contexte de ce chapitre, nous entendons par *système de transport des marchandises*, le système, éventuellement multimodal, qui permet l'acheminement de biens ou de services entre les lieux de production et de consommation. Quelques exemples illustratifs : le transport intermodal international de conteneurs, les systèmes de transport de marchandises par rail et par camion en France, les systèmes de courrier express et régulier, le système qui permet à une entreprise de vente à distance d'acheminer depuis les lieux de stockage et en moins de 24h les articles que ses clients ont commandés, celui de l'entreprise de ramassage d'ordures dans une ville ou finalement le système qui permet d'envoyer avec un délai très bref une ambulance sur les lieux d'un accident.

On distingue le transport longue distance, uni ou multimodal, et les activités de distribution ou de ramassage s'effectuant généralement dans une zone géographique relativement restreinte permettant de servir plusieurs clients dans la journée. Il faut remarquer que du point de vue des expéditeurs, les mouvements des charges requièrent, pour la plupart, des activités de ramassage et de distributions aux extrémités de la chaîne, ainsi qu'une suite de mouvements longue distance.

Nous retrouvons deux grandes types d'organisation dans le transport longue distance : des systèmes avec consolidation, où un seul véhicule ou convoi sert à déplacer les mar-

chandises de différents clients ayant des origines et des destinations finales possiblement différentes, ainsi que des transporteurs « porte-à-porte» qui personnalisent leurs services pour chaque client spécifique.

Le camionnage en charge complète constitue un exemple typique d'un transport personnalisé. Dans ce mode de transport, un véhicule est habituellement réservé à chaque client. Lorsque le client appelle, on assigne un camion et un conducteur ou une équipe de conducteurs à cette tâche. Le camion se rend au lieu désigné par le client, où il est chargé, puis il se dirige vers la destination indiquée. À destination, le camion est déchargé et le conducteur appelle le répartiteur pour demander une nouvelle affectation. Le répartiteur peut lui assigner une nouvelle charge, lui demander de se rendre à vide vers un nouveau lieu où une demande devrait bientôt se présenter ou demander au conducteur d'attendre et de rappeler plus tard.

Le transporteur en charge complète évolue donc dans un environnement extrêmement dynamique où on ne connaît pas avec certitude les demandes futures, les temps de déplacement, les périodes d'attente aux installations des clients, les positions exactes des véhicules, chargés ou vides, à des moments ultérieurs, et ainsi de suite. Le service est adapté à chaque client, et l'affectation rapide des véhicules à des demandes rentables est de la plus grande importance. L'élaboration de stratégies efficaces en matière de gestion et d'affectation des ressources est par conséquent au cœur du processus de gestion. Ces stratégies visent à maximiser le volume de demandes satisfaites (charges déplacées) et les profits qui y sont associés, tout en favorisant une utilisation optimale des ressources disponibles : conducteurs, flottes de tracteurs et de semi-remorques, etc. Les services assurés par des navires en location, ainsi que nombre de problèmes de planification opérationnelle des activités dans les systèmes avec consolidation partagent ces caractéristiques dynamiques et stochastiques. Nous traitons de ces aspects dans la section 4.4 de ce chapitre.

Le transport personnalisé n'est cependant pas toujours la réponse appropriée aux besoins de l'expéditeur. Les relations et compromis entre le volume et la fréquence des expéditions, d'une part, et entre le coût, la fréquence et le temps de livraison, d'autre part, dictent souvent l'utilisation de services de transport avec consolidation. Dans le transport de voyageurs, ce choix équivaut, par exemple, à prendre un taxi ou à utiliser les services de transport en commun. Le transport consolidé de marchandises est effectué par des transporteurs routiers en charges brisées (LTL), des chemins de fer, des lignes maritimes, des services postaux réguliers et express, etc. Dans certains pays où une autorité centrale contrôle plus ou moins une large part du système de transport, le transport des marchandises appartient aussi à cette catégorie.

Les transporteurs de marchandises consolidées sont organisés sous la forme de réseaux en étoile (« hub-and-spoke »). Dans les systèmes de ce type, on offre le service entre un nombre de points origine-destination beaucoup plus grand que le nombre de services directs du point d'origine au point de destination assurés par le transporteur. Par conséquent, et pour tirer avantage des économies d'échelle, les demandes faibles sont d'abord acheminées vers un point intermédiaire – un terminal de consolidation ou « hub

» – tel qu'un aéroport, un port, une cour de triage ou une plateforme intermodale. Au terminal, le trafic est trié et groupé – consolidé – et des flux plus volumineux sont acheminés vers d'autres terminaux par des services à grande fréquence et à grande capacité. Les services à plus faible fréquence, qui exploitent souvent des véhicules de plus petite dimension, sont utilisés entre les hubs et les terminaux d'origine et de destination.

Une telle organisation permet d'offrir un service de fréquence et qualité supérieures entre les terminaux de consolidation et d'assurer une utilisation plus efficace des ressources. L'inconvénient de ce type d'organisation réside dans les délais accrus attribuables à des itinéraires plus longs et au temps passé dans les terminaux. Cet aspect explique en partie pourquoi on ne trouve pratiquement aucun système de réseau en étoile « pur » le transport direct étant organisé pour des paires origine-destination à forte demande ou à haute priorité. L'organisation des services de transport avec consolidation demande donc des activités et des méthodes de planification plutôt complexes. Nous présentons à la section 4.1 les principales classes de modèles de recherche opérationnelle proposées à ces fins.

Les activités de ramassage et de livraison sont omniprésentes et représentent un maillon essentiel des chaînes logistiques, d'approvisionnement et de distribution à courte et longue distance des biens et des services. Sous diverses formes, on retrouve ces mêmes problèmes dans la planification de la production, la gestion des ateliers et entrepôts flexibles, la planification et la gestion des réseaux de télécommunication, et ainsi de suite. Un très grand nombre de contributions de la recherche opérationnelle a été dédié à ces problèmes et les sections 5 à 7 en présentent les principales.

Les terminaux jouent un rôle essentiel dans les systèmes de transport de marchandise. Dans le transport avec consolidation, les terminaux offrent des services de chargement/déchargement des véhicules, de formation et de séparation des convois, de tri, consolidation et transfert des charges. C'est des terminaux que partent les routes de ramassage et de livraison et c'est à ces mêmes terminaux que les véhicules retournent après avoir effectuer leurs routes. De façon plus générale, des « terminaux » constituent les divers points de rupture de charge des flux de marchandise. Le choix des emplacements des terminaux, ainsi que des points de rupture de charge, correspond à une grande famille de problèmes de recherche opérationnelle : les problèmes de localisation. Nous en présentons trois classes importantes à la section 3.

3 Conception de systèmes

Cette section est consacrée aux modèles visant à résoudre les problèmes de planification stratégique auxquels sont confrontés les transporteurs. Nous nous intéressons plus particulièrement aux décisions se rapportant à la conception du réseau d'infrastructure physique : où localiser les terminaux ? Quel type et quelle quantité d'équipement mettre

en place à chaque installation ? Quel type de lignes ou quelle capacité ajouter ? Quelles lignes ou installations abandonner ? Quels clients ou quelles régions relier par des services directs et comment desservir ceux ou celles qui ne sont pas reliés par de tels services ? Et ainsi de suite. Nous rassemblons ces problèmes sous la dénomination *conception de systèmes*.

Les modèles de conception de systèmes pour le transport des marchandises prennent généralement la forme de formulations discrètes de localisation destinées à résoudre des problèmes associés à la localisation de terminaux de consolidation et à l'acheminement de la demande entre les terminaux d'origine et de destination. Des modèles plus généraux de conception de réseaux sont proposés lorsque des décisions reliées à l'offre de service de transport font partie du processus de planification. Dans ce qui suit, nous présentons trois classes de formulations associées à la planification stratégique des systémes de transport avec consolidation et plus généralement à la planification du transport intermodal et des chaînes logistiques. Pour plus de détails sur les modèles de localisation et de conception de réseaux et les méthodes de résolution associées on peut consulter : Ahuja, Magnanti et Orlin [3], Balakrishnan, Magnanti et Mirchandani [14], Crainic [39], Crainic et Kim [48], Crainic et Laporte [49], Daskin [56], Daskin et Owen [57], Drezner [64], Labbé, Peeters et Thisse [112], Labbé et Louveaux [111], Magnanti et Wong [122], Minoux [124], Mirchandani et Francis [125], Nemhauser et Wolsey [131] et Salkin et Mathur [167].

3.1 Localisation avec exigences d'équilibrage

Le problème de localisation multiproduit avec exigences d'équilibrage a été présenté pour la première fois par Crainic, Dejax et Delorme [42] dans le contexte de la gestion d'une flotte hétérogène de conteneurs par une société internationale de transport maritime. Le déroulement des opérations terrestres du transporteur se résume comme suit : Quand un navire arrive au port, la société doit livrer des conteneurs chargés, qui peuvent être de différents types et dimensions, vers des destinations intérieures préalablement désignées. À la suite de leur déchargement par le client importateur, les conteneurs vides sont déplacés vers un dépôt. De là, les conteneurs vides peuvent être livrés ultérieurement aux clients qui en font la demande en vue d'une expédition subséquente de leurs propres produits. En outre, des conteneurs vides doivent être déplacés, *repositionnés* vers d'autres dépôts. Ces mouvements entre les dépôts résultent des déséquilibres régionaux entres les disponibilités et les besoins en conteneurs vides à la grandeur du réseau : certains secteurs manquent de conteneurs de certains types, alors que d'autres présentent des surplus. Il faut donc équilibrer les mouvements de conteneurs vides entre les dépôts, ce qui distingue ce problème des applications classiques de localisation-allocation.

Le problème général consiste à localiser les dépôts et à allouer les clients aux dépôts (pour chaque type de conteneur et direction de mouvement) afin de recueillir les conteneurs vides disponibles chez les clients importateurs (offreurs de conteneurs vides) et de satisfaire les demandes en conteneurs vides des clients exportateurs (demandeurs), tout

en minimisant les coûts d'exploitation totaux : les coûts d'ouverture et d'exploitation des dépôts ainsi que les coûts occasionnés par les mouvements client-dépôt et interdépôts. Les coûts de transport sont spécifiques à chaque type de conteneur et mode de transport. Les mouvements interdépôts sont exécutés par des modes de transport à forte densité (généralement le rail) et leur coût unitaire de transport est inférieur à celui des autres types de déplacements.

La formulation est fondée sur un réseau orienté $\mathcal{G} = (\mathcal{N}, \mathcal{A})$, où $\mathcal{N}$ représente l'ensemble des nœuds ou sommets et $\mathcal{A}$ est l'ensembles des arcs ou liens. L'ensemble des nœuds peut être subdivisé en trois sous-ensembles : $\mathcal{O}$, l'ensemble des nœuds d'origine (clients offreurs) ; $\mathcal{D}$, l'ensemble des nœuds de destination (clients demandeurs) ; et $\mathcal{H}$, l'ensemble des hubs ou dépôts. Les ensembles de clients adjacents à chaque dépôt $j \in \mathcal{H}$ sont représentés par $\mathcal{O}(j) = \{i \in \mathcal{O} : (i,j) \in A\}$ et $\mathcal{D}(j) = \{i \in \mathcal{D} : (j,i) \in A\}$. On suppose qu'il existe au moins une origine ou une destination adjacente à chaque dépôt. Les ensembles de dépôts adjacents à chaque nœud $i \in N$ dans les deux directions sont représentés par $\mathcal{H}^+(i) = \{j \in \mathcal{H} : (i,j) \in A\}$, et $\mathcal{H}^-(i) = \{j \in \mathcal{H} : (j,i) \in A\}$. Les mouvements directs de conteneurs client-client ne sont pas permis dans ce problème. Par conséquent, l'ensemble des arcs $\mathcal{A}$ peut être subdivisé en trois sous-ensembles : arcs client-dépôt, $A_{OH} = \{(i,j) \in A : i \in \mathcal{O},\ j \in \mathcal{H}\}$, arcs dépôt-client, $A_{HD} = \{(i,j) \in A : i \in \mathcal{H},\ j \in \mathcal{D}\}$ et arcs dépôt-à-dépôt, $A_{HH} = \{(i,j) \in A : i \in \mathcal{H},\ j \in \mathcal{H}\}$. Les produits (types de conteneurs) qui circulent sur le réseau sont représentés par l'ensemble $\mathcal{P}$.

Pour chaque client offreur, $i \in \mathcal{O}$, l'offre du produit p est exprimée sous la forme $o_i^p \geq 0$, tandis que pour chaque client demandeur, $i \in \mathcal{D}$, la demande du produit p est représentée par $d_k^p \geq 0$. Un coût non négatif c_{ij}^p est engagé pour chaque unité de produit p qui se déplace sur un arc (i,j). En outre, pour chaque dépôt $j \in \mathcal{H}$, un coût non négatif fixe f_j est engagé si on utilise (ouvre) le dépôt. Soit x_{ij}^p la quantité de produit p se déplaçant sur un arc (i,j) et y_j, la variable binaire de localisation qui prend la valeur 1 si le dépôt j est ouvert, et 0 sinon. On formule alors le problème comme suit :

$$\text{Minimiser} \quad \sum_{j \in \mathcal{H}} f_j y_j + \sum_{p \in \mathcal{P}} \left(\sum_{(i,j) \in A_{OH}} c_{ij}^p x_{ij}^p + \sum_{(j,i) \in A_{HD}} c_{ji}^p x_{ji}^p + \sum_{(j,k) \in A_{HH}} c_{jk}^p x_{jk}^p \right) \tag{1}$$

$$\text{sujet à} \quad \sum_{j \in \mathcal{H}^+(i)} x_{ij}^p = o_i^p \quad i \in \mathcal{O},\ p \in \mathcal{P}, \tag{2}$$

$$\sum_{j \in \mathcal{H}^-(i)} x_{ji}^p = d_k^p \quad i \in \mathcal{D},\ p \in \mathcal{P}, \tag{3}$$

$$\sum_{i \in \mathcal{D}(j)} x_{ji}^p + \sum_{k \in \mathcal{H}^+(j)} x_{jk}^p - \sum_{i \in \mathcal{O}(j)} x_{ij}^p - \sum_{k \in \mathcal{H}^-(j)} x_{kj}^p = 0 \quad j \in \mathcal{H},\ p \in \mathcal{P}, \tag{4}$$

$$x_{ij}^p \leq o_i^p y_j \quad j \in \mathcal{H}, i \in \mathcal{O}(j), p \in \mathcal{P}, \tag{5}$$

$$x_{ji}^p \leq d_k^p y_j \quad j \in \mathcal{H}, i \in \mathcal{D}(j), p \in \mathcal{P}, \tag{6}$$

$$x_{ij}^p \geq 0 \qquad (i,j) \in A, p \in \mathcal{P}, \tag{7}$$

$$y_j \in \{0,1\} \qquad j \in \mathcal{H}. \tag{8}$$

Les contraintes (2) et (3) garantissent que les exigences d'offre et de demande sont respectées, les relations (4) correspondent aux contraintes de conservation des flux aux dépôts, tandis que les équations (5) et (6) interdisent les mouvements entre clients et les dépôts fermés. Les contraintes analogues pour les flux interdépôts sont redondantes si les coûts entre les dépôts satisfont l'inégalité du triangle et ne sont pas incluses dans la formulation.

Le problème a été résolu à l'aide d'heuristiques d'ascension duale (Crainic et Delorme [43]), d'algorithmes branch-and-bound séquentiels et parallèles (Crainic, Delorme et Dejax [44] ; Gendron et Crainic [77, 78] ; Bourbeau, Gendron et Crainic [20]) et de recherche avec tabous séquentielle et parallèle (Crainic *et al.* [47] ; Crainic, Toulouse et Gendreau [54, 53, 52] ; Gendron, Potvin et Soriano [79]). Gendron, Potvin et Soriano [81, 80]) ont étudié la version du problème avec capacités et proposé des métaheuristiques séquentielles et parallèles.

3.2 Production-distribution multiproduit

Le problème de production-distribution multiproduit est une version simplifiée du cas précédent où aucun mouvement entre des hubs n'est pris en compte. Ici, les produits peuvent être acheminés depuis leurs origines $i \in \mathcal{O}$ jusqu'à leurs destinations $k \in \mathcal{D}$ soit directement, soit en passant par un terminal de consolidation $j \in \mathcal{H}$. Les ensembles $\mathcal{O}, \mathcal{D}$ et $\mathcal{H}$ ne sont pas nécessairement disjoints. Les décisions incluses dans la formulation sont le nombre et la localisation des terminaux de consolidation de même que la structure des flux de produits à travers le système : soit directement, soit par un terminal de consolidation.

La capacité d'un terminal de consolidation localisé au site $j \in \mathcal{H}$ est notée u_j. Le coût de transport par unité de produit p depuis l'origine i jusqu'à la destination k transitant par le hub j est noté c_{ijk}^p, alors que c_{ik}^p représente le coût unitaire de transport lorsque des mouvements directs sont autorisés. Les variables de décision d'acheminement des flux correspondantes sont désignées par x_{ijk}^p et x_{ik}^p, respectivement. Toute autre notation est la même que dans le problème précédent. Le modèle prend la forme suivante :

$$\text{Minimiser} \quad \sum_{j \in \mathcal{H}} f_j y_j + \sum_{p \in \mathcal{P}} \sum_{i \in \mathcal{O}} \sum_{k \in \mathcal{D}} \left(c_{ik}^p x_{ik}^p + \sum_{j \in \mathcal{H}} c_{ijk}^p x_{ijk}^p \right) \tag{9}$$

$$\text{sujet à} \quad x_{ijk}^p \leq min\{o_i^p, u_j, d_k^p\} y_j \qquad i \in \mathcal{O},\ j \in \mathcal{H}\ , k \in \mathcal{D},\ p \in \mathcal{P} \tag{10}$$

$$\sum_{p \in \mathcal{P}} \sum_{i \in \mathcal{O}} \sum_{k \in \mathcal{D}} x_{ijk}^p \leq u_j y_j \qquad j \in \mathcal{H} \tag{11}$$

$$\sum_{k \in \mathcal{D}} \left(x_{ik}^p + \sum_{j \in \mathcal{H}} x_{ijk}^p \right) \leq o_i^p \qquad i \in \mathcal{O},\ p \in \mathcal{P} \tag{12}$$

$$\sum_{i \in \mathcal{O}} \left(x_{ik}^p + \sum_{j \in \mathcal{H}} x_{ijk}^p \right) = d_k^p \qquad k \in \mathcal{D},\ p \in \mathcal{P} \tag{13}$$

$$y_j \in \{0, 1\} \qquad j \in \mathcal{H} \tag{14}$$

$$x_{ijk}^p \geq 0\ , x_{ik}^p \geq 0 \qquad i \in \mathcal{O}\ , k \in \mathcal{D},\ j \in \mathcal{H},\ p \in \mathcal{P}. \tag{15}$$

Dans cette formulation, l'équation (9) calcule le coût total de l'ouverture des hubs et de la distribution des produits. Les contraintes (10) et (11) assurent que seuls les terminaux ouverts sont utilisés. La relation (11) force aussi le respect de la contrainte de capacité des terminaux de consolidation. Les offres et les demandes de marchandises sont satisfaites par les relations (12) et (13), respectivement.

Lorsque la demande est définie par rapport aux paires origine-estination, on modifie la formulation précédente en remplaçant les relations (10) par (16) et les contraintes (12) et (13) par (17) :

$$x_{ijk}^p \leq min\{u_j, d_{ik}^p\} y_j \quad i \in \mathcal{O},\ j \in \mathcal{H}\ , k \in \mathcal{D},\ p \in \mathcal{P} \tag{16}$$

$$x_{ik}^p + \sum_{j \in \mathcal{H}} x_{ijk}^p = d_{ik}^p \qquad i \in \mathcal{O},\ k \in \mathcal{D},\ p \in \mathcal{P}. \tag{17}$$

où d_{ik}^p représente la quantité de produit p à transporter depuis le terminal d'origine i jusqu'au terminal de destination k. Il existe de nombreuses variantes de ces formulations provenant, par exemple, de la spécification de coûts ou des capacités partielles associés à des produits spécifiques (voir Aikens [4], par exemple). Toutes ces formulations sont difficiles et des heuristiques sont généralement utilisées (Erlenkotter [67] ; Pirkul et Jayaraman [139, 140] ; ainsi que les références indiquées au début de la sous-section).

3.3 Localisation de hubs

Les modèles de localisation de hubs s'appliquent lorsqu'il faut déterminer simultanément la localisation des terminaux de consolidation et l'allocation des terminaux régionaux ou des clients à ces derniers. Le modèle de base assume qu'il n'y a ni capacités aux hubs, ni coûts « d'ouverture » pour les liens entre les terminaux régionaux et les terminaux de consolidation, que le transport direct n'est pas permis entre les terminaux régionaux et que tout le trafic d'une origine à une destination passe par deux hubs. Comme c'est le cas pour tous les problèmes fondés sur des hubs, on suppose que

le transport entre les hubs est plus efficace en raison de la concentration des flux, ce qui entraîne un coût unitaire inférieur à ceux des autres mouvements dans le système. La notation est dans une large mesure similaire à celle des modèles précédents. S'y ajoutent les définitions suivantes :

$y_j = 1$, si un hub est localisé au site j, 0 sinon ;

$y_{ij} = 1$, si le terminal i est relié au hub j, 0 sinon ;

x^p_{ijlk} : quantité de flux du produit p avec une origine i et une destination k qui circule à travers les hubs j et l, dans cet ordre.

La formulation suivante suppose que P hubs doivent être localisés parmi les $|\mathcal{H}|$ sites potentiels :

$$\text{Minimiser} \quad \sum_{p\in\mathcal{P}}\left\{\sum_{i\in\mathcal{O}}\sum_{j\in\mathcal{H}} c^p_{ij}y_{ij}\left(\sum_{l\in\mathcal{H}}\sum_{k\in\mathcal{D}} x^p_{ijlk}\right) + \sum_{l\in\mathcal{H}}\sum_{k\in\mathcal{D}} c^p_{lk}y_{lk}\left(\sum_{i\in\mathcal{O}}\sum_{j\in\mathcal{H}} x^p_{ijlk}\right) + \sum_{j\in\mathcal{H}}\sum_{l\in\mathcal{H}} c^p_{jl}y_{ij}y_{lk}\left(\sum_{i\in\mathcal{O}}\sum_{k\in\mathcal{D}} x^p_{ijlk}\right)\right\} \tag{18}$$

$$\text{sujet à} \quad \sum_{j\in\mathcal{H}} y_j = P \tag{19}$$

$$\sum_{j\in\mathcal{H}}\sum_{l\in\mathcal{H}} x^p_{ijlk} = d^p_{ik} \qquad i\in\mathcal{O},\ k\in\mathcal{D},\ p\in\mathcal{P}, \tag{20}$$

$$y_{ij} \le y_j \qquad i\in\mathcal{O},\ j\in\mathcal{H}, \tag{21}$$

$$y_{lk} \le y_l \qquad k\in\mathcal{D},\ l\in\mathcal{H}, \tag{22}$$

$$x^p_{ijlk} \le d^p_{ik}y_j \qquad i\in\mathcal{O},\ k\in\mathcal{D},\ j,l\in\mathcal{H},\ p\in\mathcal{P}, \tag{23}$$

$$x^p_{ijlk} \le d^p_{ik}y_l \qquad i\in\mathcal{O},\ k\in\mathcal{D},\ j,l\in\mathcal{H},\ p\in\mathcal{P}, \tag{24}$$

$$y_j = 0 \text{ ou } 1 \qquad j\in\mathcal{H}, \tag{25}$$

$$y_{ij} = 0 \text{ ou } 1 \qquad i\in\mathcal{O},\ j\in\mathcal{H}, \tag{26}$$

$$x^p_{ijlk} \ge 0 \qquad i\in\mathcal{O},\ k\in\mathcal{D},\ j,l\in\mathcal{H},\ p\in\mathcal{P}. \tag{27}$$

La fonction objectif (18) minimise le coût de transport total du système. En raison de sa similitude avec le problème de localisation *p-médiane*, cette classe de formulations porte le nom de problème de *p-hub médiane* (Campbell [26]). Il faut souligner, cependant, qu'en dépit de ces similitudes de nom, plusieurs propriétés des problèmes de p-médiane ne s'appliquent pas aux problèmes p-hub (ainsi, par exemple, l'affectation des nœuds de demande à l'installation ouverte la plus proche n'est pas optimale pour les problèmes de p-hub médiane).

La formulation (18) - (27) est quadratique et elle a été présentée pour la première fois par O'Kelly [132]. Campbell [26] a introduit le premier mécanisme de linéarisation, perfectionné ensuite par Campbell [27], Skorin-Kapov, Skorin-Kapov et O'Kelly [170] et

O'Kelly *et al.* [135]. Le processus produit une formulation linéaire en nombres entiers mixtes fondée sur les *chemins* :

$$\text{Minimiser} \quad \sum_{p\in\mathcal{P}}\sum_{i\in\mathcal{O}}\sum_{j\in\mathcal{H}}\sum_{l\in\mathcal{H}}\sum_{k\in\mathcal{D}} c^p_{ijlk}x^p_{ijlk} \tag{28}$$

sujet à (19), (20), (23) à (25) et (27), où c^p_{ijlk} représente le coût unitaire de transport pour le produit p de l'origine i à la destination k, avec consolidation aux hubs j et l, dans cet ordre.

Les deux formulations précédentes permettent des affectations multiples des terminaux d'origine et de destination aux hubs. Pour forcer une affectation unique des terminaux aux hubs, on remplace les contraintes (23) et (24) de la première formulation par

$$\sum_{j\in\mathcal{H}} y_{ij} = 1 \qquad i \in \mathcal{O}, \tag{29}$$

$$\sum_{l\in\mathcal{H}} y_{lk} = 1 \qquad k \in \mathcal{D}, \tag{30}$$

$$x^p_{ijlk}/d^p_{ik} = y_{ij} \quad i \in \mathcal{O},\ k \in \mathcal{D},\ j,l \in \mathcal{H},\ p \in \mathcal{P}, \tag{31}$$

$$x^p_{ijlk}/d^p_{ik} = y_{lk} \quad i \in \mathcal{O},\ k \in \mathcal{D},\ j,l \in \mathcal{H},\ p \in \mathcal{P}, \tag{32}$$

où les variables y_{ij} et y_{lk} correspondent aux décisions d'aller ou non directement du terminal i au hub j et du hub l au terminal k, respectivement.

Ces problèmes sont difficiles à résoudre. Par conséquent, les méthodes de résolution proposées sont essentiellement des heuristiques (O'Kelly [132] ; Aykin [9] ; Klincewicz [106] ; Campbell [27] ; etc.). Des mét-aheuristiques ont amélioré ces résultats : réseaux neuronaux (Smith, Krishnamoorthy et Palaniswami [171]), recuit simulé (Ernst et Krishnamoorthy [68]) et recherches avec tabous (Klincewicz [107] ; Skorin-Kapov et Skorin-Kapov [169]). Des bornes inférieures ont aussi été proposées (O'Kelly [133] ; O'Kelly, Skorin-Kapov et Skorin-Kapov [136] ; Skorin-Kapov, Skorin-Kapov et O'Kelly [170]), ainsi qu'un certain nombre de méthodes qui produisent la solution optimale pour des problèmes de taille limitée (Skorin-Kapov, Skorin-Kapov et O'Kelly [170] ; O'Kelly *et al.* [135] ; Ernst et Krishnamoorthy [68]).

On peut formuler un certain nombre d'extensions à ces modèles. On peut ainsi ajouter des termes de coût fixe $\sum_{j\in\mathcal{H}} f_j y_j$ à la fonction objectif (et abandonner la contrainte (19) sur le nombre de hubs à ouvrir) pour représenter les coûts d'ouverture des centres de consolidation (O'Kelly [134]). Klincewicz [108] a proposé des heuristiques duales pour cette formulation, alors qu'Abdinnour-Helm et Venkataramanan [2] ont mis au point une métaheuristique génétique de recherche et un algorithme de séparation et évaluation progressive (branch-and-bound). La taille des problèmes étudiés été limitée. D'autres extensions prennent en compte des coûts fixes pour l'allocation de terminaux d'origine et de destination aux hubs, des capacités aux hubs, des schémas d'acheminement plus

complexes (comportant plus que deux hubs, par exemple), des niveaux minimaux de trafic requis afin de permettre une connexion terminal-hub, etc. Ces formulations de conception des réseaux reflètent ainsi mieux la complexité des systèmes de transport mais sont très difficiles à résoudre, et peu de pistes ont été proposées jusqu'à ce jour (Aykin [10, 12, 11] ; Kuby et Gray [110] ; Jaillet, Song, Yu [102]). Il est essentiel de poursuivre la recherche dans ce domaine.

4 Transport longue distance

Nous examinons des modèles proposés pour deux grandes classes de probleèmes de planification et de gestion du transport à longue distance. La première classe de problème concerne la conception des réasux de services des transporteurs avec consolidation. Ces problèmes se retrouvent généralement au niveaux stratégiques/tactiques de planification des opérations de telles entreprises. La seconde classe regroupe des problèmes de gestion de flotte, des mouvements de véhicules vides, en particulier, correspondant au niveau opérationnel de gestion du système.

4.1 Conception de réseaux de service

Cette section est consacrée aux modèles dédiés aux problèmes de planification tactique auxquels sont confrontés les transporteurs avec consolidation : sur quels parcours doit-on fournir le service ? Quel type de service (mode) doit-on utiliser ? À quelle fréquence doit-on offrir le service sur chaque parcours et selon quel horaire ? Comment acheminer les charges à travers les réseaux physiques et de service ? Comment répartir le travail entre les terminaux du système ?

Les formulations de conception de réseaux de service visent à bâtir un plan de transport (ou plan de chargement) afin d'assurer le fonctionnement efficace du système, la satisfaction du client et la rentabilité de l'entreprise. L'élément « offre » de cette équation suppose une représentation intégrée des opérations sur l'ensemble du réseau, incluant les activités de consolidation dans les terminaux, ainsi que la sélection, la composition, l'ordonnancement et l'acheminement des services. Du côté de la « demande », l'acheminement des marchandises à travers le réseau doit être planifié afin d'assurer une livraison rapide et fiable en conformité avec les spécifications du client et les objectifs propres au transporteur.

Les objectifs du processus sont aussi complexes. Traditionnellement, les attentes des clients se résumaient au souhait de voir les marchandises être livrées au plus bas coût possible. Cette demande, conjuguée à l'attention que porte habituellement toute entreprise à ses coûts, explique pourquoi le premier objectif d'un transporteur de marchandises

a jusqu'à maintenant été – et est encore aujourd'hui pour beaucoup de transporteurs – d'exercer ses activités au plus bas coût possible. Néanmoins, on constate de plus en plus que les clients ne recherchent pas seulement des tarifs bas, ils exigent aussi un service de qualité, principalement sur les plans de la vitesse, de la flexibilité et de la fiabilité. La forte augmentation de la part de marché occupée par les transporteurs routiers est largement attribuable à ce phénomène. Par conséquent, l'un des principaux buts de la planification tactique, consiste à trouver le meilleur équilibre entre les coûts d'exploitation, la rentabilité de l'entreprise et la qualité des services que l'on mesure, dans la plupart des cas, aux retards des marchandises et des véhicules ou au respect des objectifs de délais de livraison prédéterminés.

Les problèmes de conception de réseaux de service portent donc sur deux principaux types de décisions. Le premier consiste à déterminer le réseau de service, c'est-à-dire à sélectionner les parcours – les terminaux d'origine et de destination, la route physique et les arrêts intermédiaires – sur lesquels les services seront offerts ainsi que les caractéristiques de chaque service : mode, fréquence ou horaire, etc. Le second type de décisions vise à déterminer la distribution du trafic, c'est-à-dire, les itinéraires (parcours) utilisés pour déplacer le flux de chaque demande : services utilisés, terminaux par lesquels passe la marchandise, opérations exécutées dans ces terminaux, etc. Le réseau de service précise les mouvements, dans l'espace et dans le temps, des véhicules et des convois des divers modes pris en compte. Les itinéraires déterminent les flux des marchandises sur les services et à travers les terminaux du réseau de service.

On peut établir une distinction entre les formulations de conception de réseaux de service statiques et dynamiques. Dans la première catégorie, on assume que la demande ne varie pas durant la période de planification visée. La dimension temporelle du réseau de service est alors considérée de façon implicite dans la définition des services et des opérations aux terminaux. La dimension temporelle des processus décisionnels est explicitement prise en compte dans les formulations dynamiques. Celles-ci incluent une représentation explicite des mouvements dans le temps et ciblent habituellement la planification des horaires pour soutenir les décisions liées au moment du départ des services. La plupart des modèles de conception de réseaux de service proposés dans la littérature prennent la forme de formulations de conception de réseaux déterministes multiproduit avec coûts fixes et capacités. Pour des sythèses des contributions au domaine, on peut consulter Assad [6], Cordeau, Toth et Vigo [36], Christiansen *et al.* [31], Crainic [38, 39, 40], Crainic et Laporte [49], Crainic et Kim [48] et Delorme, Roy et Rousseau [60].

4.2 Formulations statiques

Dans le cas statique, la planification du réseau de service est basée sur un réseau $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ sur lequel le transporteur peut déplacer ses véhicules. Dans ce réseau, le transporteur exploite (ou utilise) un certain nombre de terminaux, dont les caractéristiques

(dimensions, capacités, coûts) et fonctionnalités (consolidation ou non) varient. Les sommets dans $\mathcal{H} \subseteq \mathcal{N}$ correspondent aux nœuds où sont situés les terminaux ; on suppose, aux fins de simplification, que tous les terminaux peuvent exécuter toutes les opérations.

Le réseau de service spécifie les services de transport qui pourraient être offerts sur ce réseau. Chaque service $s \in \mathcal{S}$ se caractérise par 1) son parcours, défini comme un chemin dans $\mathcal{A}$, de son terminal d'origine à son terminal de destination, avec, éventuellement, des terminaux intermédiaires où s'arrête le service et où des tâches peuvent être exécutées (par exemple, une ligne maritime qui arrête à des ports prédéterminés pour décharger et charger des conteneurs) ; 2) sa capacité ; 3) sa classe de services qui indique des caractéristiques telles que le mode, le type des produits qui peuvent être chargées, la vitesse et la priorité, etc. La conception du réseau de service revient alors à décider quel service inclure dans le plan de transport de façon à répondre à la demande et aux objectifs du transporteur. Lorsqu'un service peut être assuré à répétition durant la période de planification (par exemple, plusieurs navires semblables partent durant le mois et visitent les mêmes ports dans le même ordre), le processus doit aussi déterminer la fréquence de chaque service envisagé.

Crainic et Rousseau [50] ont proposé une formulation multimode, multiproduit basée sur les chemins pour déterminer les fréquences des services et la distribution correspondante de la demande. Dans leur modèle, le trafic circule en empruntant des itinéraires. Un itinéraire $l \in \mathcal{L}^p$ pour un produit $p \in \mathcal{P}$ précise le chemin utilisé dans le réseau des services pour déplacer (une partie de) la demande correspondante, incluant les opérations (consolidation, transfert, arrêt) a exécuter dans les terminaux intermédiaires. La demande en produit p est notée d_p. Les décisions de distribution des flux sont représentées par les variables h_l^p qui indiquent le volume de produit p déplacé en utilisant l'itinéraire $l \in \mathcal{L}^p$.

Les fréquences de service y_s, $s \in \mathcal{S}$, définissent le niveau de service offert, c'est-à-dire à quelle fréquence chaque service est opéré durant la période de planification. Soit $y = \{y_s\}$ et $h = \{h_l^p\}$ qui représentent les vecteurs des variables de décision. Le modèle minimise le coût généralisé total du système, tout en satisfaisant la demande de transport et les normes de service :

$$\text{Minimiser} \quad \sum_{s \in \mathcal{S}} \mathcal{F}_s(y) + \sum_{p \in \mathcal{P}} \sum_{l \in \mathcal{L}^p} \mathcal{C}_l^p(y,h) + \Theta(y,h) \tag{33}$$

$$\text{sujet à} \quad \sum_{l \in \mathcal{L}^p} h_l^p = d_p \qquad p \in \mathcal{P} \tag{34}$$

$$y_s \geq 0 \text{ et entier} \qquad s \in \mathcal{S} \tag{35}$$

$$h_l^p \geq 0 \qquad l \in \mathcal{L},\ p \in \mathcal{P} \tag{36}$$

où

$\mathcal{F}_s(y)$: coût total de l'exploitation du service s ;

$\mathcal{C}_l^p(y,h)$: coût total pour déplacer (une partie de) la demande du produit p en utilisant l'itinéraire l ;

$\Theta(y, h)$: fonction de pénalité prenant en compte diverses relations et restrictions, telles que la capacité des infrastructures et services.

La fonction objectif définit le coût total du système et inclut le coût total d'exploitation d'un réseau de service à des fréquences données, le coût total de déplacement du fret en utilisant les itinéraires sélectionnés pour chaque produit, ainsi qu'un certain nombre de paramètres qui transposent les contraintes de fonctionnement et de service en valeurs monétaires. Ce troisième élément indique, par exemple, que les contraintes en matière de capacité de service sont considérées comme des objectifs d'utilisation et que la suraffectation du trafic est permise au prix de coûts et de délais supplémentaires. On peut alors quantifier les compromis entre les coûts et niveaux de service des diverses activités et identifier ainsi la meilleure solution du point de vue de l'ensemble du système.

La fonction objectif calcule un coût généralisé qui peut inclure diverses mesures de productivité associées aux activités de transport et aux opérations dans les terminaux. Outre les coûts réels engendrés par l'exécution des opérations, on pourra ainsi prendre en considération de façon explicite les coûts, délais et autres mesures de la performance liées à la qualité et à la fiabilité du service offert, évaluer des solutions de rechange et déterminer les compromis les plus avantageux. Par exemple, on a ainsi utilisé des fonctions non linéaires de congestion pour tenir compte des délais de plus en plus importants observés lorsque les installations de capacité limitée (les terminaux de consolidation, par exemple) doivent desservir un volume croissant de trafic. La notation vise à refléter cette utilisation de formes fonctionnelles plus générales.

Le modèle a été adapté et appliqué au transport ferroviaire, au camionnage en charges partielles et aux services de courrier express (Crainic, Ferland et Rousseau [45], Roy [162], Delorme et Roy [164], Crainic et Roy [51] et Roy et Crainic [163]). La méthode de résolution originale décrite par Crainic et Rousseau [50] combine une heuristique qui, à partir de valeurs initiales élevées, réduit itérativement les fréquences, et une technique d'optimisation convexe avec génération de colonnes pour distribuer le fret sur le réseau. Des métaheuristiques récentes pourraient avantageusement remplacer le mécanisme initial de modification des fréquences (les méthodes de recherche avec tabous de Ghamlouche, Crainic et Gendreau [82, 83], par exemple).

Le modèle précédent consiste en une formulation basée sur les chemins. D'autres modèles de conception de réseaux de service proposés dans la littérature utilisent des formulations fondées sur les arcs et sont axés sur les décisions de sélection des services. Ces approches donnent lieu à des formulations de conception de réseaux $\{0, 1\}$ de grande taille avec contraintes additionnelles. Nous illustrons cette tendance à l'aide de deux cas.

Le premier provient du problème de conception des services de camionnage en charges brisées. Nous suivons la formulation introduite par Powell ([142] ; voir aussi Powell et Sheffi [151, 152, 153] et Braklow *et al.* [21]). Dans cette approche, des seuils minimaux sont imposés à certains services sur le volume transporté afin que leur sélection dans le plan de transport final puisse être considérée. Les fréquences des services deviennent ainsi non plus des décisions dans la formulation, mais un résultat de l'optimisation du choix

des services et des acheminements.

Le modèle est défini à partir d'un réseau de service $\mathcal{G} = (\mathcal{T}, \mathcal{S})$ dans lequel tous les nœuds représentent des terminaux, tandis que les liens représentent des services potentiels entre ces derniers. Deux types de terminaux sont inclus : terminaux régionaux (« end-of-lines ») d'où partent et où arrivent les flux de marchandises, et les terminaux de consolidation (« breakbulk »). Les services reliant des terminaux régionaux sont très rares, particulièrement lorsque les systèmes sont de très grande taille. Par conséquence, les seules décisions de sélection des services traitées dans le modèles correspondent à des mouvements entre les terminaux de consolidation et entre ceux-ci et les terminaux régionaux. Les principaux paramètres et variabes du modèle sont :

C_{ij} : coût unitaire de transport pour un véhicule, plein ou vide, du terminal i au terminal j ;

C_i^B : coût unitaire de manutention pour un camion au terminal de consolidation i (0, aux terminaux régionaux) ;

$C_i^E(\cdot)$: fonction calculant le coût unitaire de manutention pour un camion au terminal régional i selon le nombre de services directes qui partent de i (0, aux terminaux de consolidation) ;

w_{od} : nombre de camions qui partent du terminal o avec comme destination le terminal d ;

$\mathcal{L}$: ensemble des routes permises (respectent des contraintes particulières relativement, par exemple, aux associations de terminaux régionaux aux terminaux de consolidation) ;

y_{ij} : variable de décision de sélection des services : $y_{ij} = 1$ si le service direct du terminal i au terminal j est offert, 0 autrement ;

x_{ij}^d : volume de trafic LTL sur le lien (i, j) avec destination le terminal d ; $x_{ij} = \sum_d x_{ij}^d$;

r_{ij}^d : variable auxiliaire utilisée pour simplifier la présentation des contraintes d'acheminement ;

v_{ij} : flux de camions vides de i à j ;

x_i^B : volume total de trafic LTL traité au terminal de consolidation i, c'est à dire, le trafic qui origine à i plus le trafic qui est transféré au terminal ;

M_{ij} : fréquence minimum requise afin qu'un service direct puisse être offert à partir du terminal i au terminal j ;

$F_{ij}(x_{ij})$: fréquence de service, c'est à dire le nombre de camions qui partiront pendant la période considérée du terminal i au terminal j, où,

$$F_{ij}(x_{ij}) = \begin{cases} \max\{M_{ij},\ x_{ij}\} & \text{si } x_{ij} \geq 0 \\ 0 & \text{sinon.} \end{cases} \tag{37}$$

Le modèle s'écrit

$$\text{Minimiser} \quad \sum_{(ij)\in\mathcal{S}} C_{ij}[F_{ij}(x_{ij})y_{ij} + v_{ij}] + \sum_{i\in\mathcal{T}} [C_i^B x_i^B + C_i^E(y)w_i] \tag{38}$$

$$\text{sujet à} \quad \sum_{j\in\mathcal{T}} r_{ij}^d = 1 \text{ and } \{r_{ij}^d\} \in \mathcal{L}, \quad i,j,d \in \mathcal{T} \tag{39}$$

$$x_{ij}^d = \left[w_{id} + \sum_{k\in\mathcal{T}} x_{ki}^d \right] r_{ij}^d, \quad i,j,d \in \mathcal{T} \tag{40}$$

$$r_{ij}^d \leq y_{ij}, \quad i,j,d \in \mathcal{T} \tag{41}$$

$$\sum_{j\in\mathcal{T}} v_{ij} - \sum_{k\in\mathcal{T}} v_{ki} = w_i, \quad i \in \mathcal{T} \tag{42}$$

$$w_i = \sum_{k\in\mathcal{T}} F_{ki}(x_{ki}) - \sum_{j\in\mathcal{T}} F_{ij}(x_{ij}), \quad i \in \mathcal{T} \tag{43}$$

$$y_{ij} \in \{0,1\}, \quad (i,j) \in \mathcal{S} \tag{44}$$

$$r_{ij}^d \in \{0,1\}, \quad (i,j) \in \mathcal{S} \tag{45}$$

$$x_{ij}^d,\ v_{ij}^d,\ \hat{x}_{ij} \geq 0, \quad i,j,d \in \mathcal{T} \tag{46}$$

L'objectif (38) calcule le coût total du système. Les contraintes (39) et (40) forcent le respect des règles d'acheminement et la satisfaction de la demande. La relation (41) représente la contrainte habituelle de réalisabilité qui assure que les flux de marchandises utilisent uniquement des services sélectionnés. Les équations (42) et (43) balancent les flux des véhicules vides. Les auteurs ont proposé une approche heuristique dans laquelle le réseau de service est modifié en ajoutant ou en fermant un arc de service à chaque itération, après quoi les problèmes d'acheminement des charges et de balancement des véhicules sont réoptimisés.

La conception de services postaux constitue le deuxième cas illustrant les approches basées sur les flux d'arc. Ces applications englobent le transport aérien et terrestre de longue distance ainsi que des activités locales de ramassage et de livraison et donnent lieu à des formulations de conception de réseaux de service complexes et de grande taille.

Le transport terrestre de longue distance est souvent assuré par camion, mais on utilise aussi le rail dans certains pays, particulièrement en Europe. À cet égard, il faut souligner que l'expansion de réseaux ferroviaires à grande vitesse rend le rail de plus en plus intéressant pour les services express. Le mode aérien est cher, mais efficace pour le transport de charges relativement volumineuses sur de longues distances. L'élément « coût » de cette équation donne lieu à des efforts particuliers qui visent à optimiser la capacité disponible et à réduire le nombre d'avions utilisés. Cette situation se traduit par un nombre relativement important d'options en matière de types d'avions, de capacités et de coûts, ce qui ajoute à la complexité du problème. Kuby et Gray [110] ont mis au point un premier modèle pour la conception du réseau d'une entreprise de livraison de colis express. Il s'agit d'un modèle de conception de réseau basé sur des chemins $\{0,1\}$, où il faut sélectionner des parcours d'avion à escales multiples en provenance et à destination d'un hub donné. Des chemins ont été générés *a priori* et le modèle a été résolu au moyen d'un programme en nombres entiers mixtes standard.

Kim *et al.* [105] et Armacost, Barnhart et Ware [5] proposent une approche plus globale pour la conception de la version multimodale du problème, où sont considérés des véhicules aériens et terrestres (Barnhart et Schneur [16] présentent une version simplifiée du problème à un seul hub). Dans ce problème, on prend en considération plusieurs hubs et types d'avions, tandis que les camions peuvent exécuter des activités de ramassage et de livraison ainsi que du transport sur des distances limitées. On s'intéresse plus spécialement à la conception de services aériens visant à assurer une livraison le lendemain. Un seul produit – les colis – est considéré dans l'application, mais il faut distinguer les flux par origine et par destination.

Le modèle prend en compte deux types de terminaux : les points d'accès par lesquels les colis entrent dans le réseau et en sortent, et les hubs, où les colis sont déchargés des avions en arrivage, triés puis chargés sur un avion en partance pour être livrés à leurs points de destination. L'avion peut circuler sans faire d'escale entre un point d'accès et un hub ou arrêter à un point d'accès en cours de route. Les restrictions de fenêtres de temps relativement aux horaires de ramassage et de livraison aux points d'accès et aux périodes de tri aux hubs limitent le nombre d'escales et imposent des contraintes sur la conception du réseau. Des avions de différents types sont disponibles en nombre limité. À chaque type d'avion correspondent un coût d'exploitation et une capacité ainsi que des caractéristiques de fonctionnement (type de moteur, autonomie, vitesse, etc.) qui déterminent les parcours qu'il peut emprunter. Le but consiste à concevoir l'ensemble d'itinéraires de coût minimum, les affectations d'avions pour ces itinéraires et les flux de colis qui répondent aux objectifs en matière de demande et de niveaux de service, tout en respectant les caractéristiques de fonctionnement des terminaux (la capacité, par exemple), le type d'avion (capacité, autonomie, vitesse) et la flotte d'avions (nombre d'avions disponibles).

Notons $\mathcal{H}$ l'ensemble de hubs et $\{\mathcal{O}(p), \mathcal{D}(p),\ p \in \mathcal{P}\}$ l'ensemble de points d'accès. Chaque produit $p \in \mathcal{P}$ a une demande d^p qui doit être déplacée de son origine $\mathcal{O}(p)$ à sa destination $\mathcal{D}(p)$. Soit $\mathcal{F}$ qui représente l'ensemble de types d'avions disponibles et $\mathcal{R}^f$ qui correspond à l'ensemble de parcours (la séquence de points d'accès qui commence ou finit à un hub) que peut emprunter un avion de type $f \in \mathcal{F}$. Les parcours deviennent les arcs de vol $\mathcal{A}$ qui relient le hub et les nœuds des points d'accès $\mathcal{N}$ du système. On note le coût du parcours de vol $r \in \mathcal{R}^f$ avec un avion de type $f \in \mathcal{F}$ par c_r^f, tandis que u_r^f représente sa capacité. Les coûts associés aux flux de colis sont négligeables par rapport à ceux liés à l'exploitation de la flotte aérienne et c'est pourquoi ils ne sont pas inclus dans la formulation. Le nombre d'avions de type f disponibles correspond à n_f et la capacité d'atterrissage du hub $h \in \mathcal{H}$ est a_h. Trois indicateurs sont également définis : 1) $\delta_{ij}^{rf} = 1$ quand l'arc de vol (i, j) est inclus dans le parcours r emprunté par un avion de type f et 0 sinon ; 2) $\beta_i^r = 1$ (-1) si le nœud i est le point d'origine (de destination) du parcours r et 0 sinon ; et 3) $\delta_r^h = 1$ si le hub h est inclus dans le parcours r et 0 sinon.

Deux types de variables de décision sont définies : y_r^f indique le nombre de fois que le parcours $r \in \mathcal{R}^f$ est emprunté par un avion de type $f \in \mathcal{F}$ et x_{ij}^p représente la quantité de produit p sur le lien aérien $(i, j) \in \mathcal{A}$. La formulation du modèle de conception de

réseaux de fréquence de desserte de base devient alors :

$$\text{Minimiser} \quad \sum_{f\in\mathcal{F}}\sum_{r\in\mathcal{R}^f} c_r^f y_r^f \tag{47}$$

$$\text{sujet à} \quad \sum_{j\in\mathcal{N}} x_{ij}^p - \sum_{j\in\mathcal{N}} x_{ji}^p = \begin{cases} d^p & \text{si } i=\mathcal{O}(p), \\ -d^p & \text{si } i=\mathcal{D}(p), \\ 0 & \text{sinon,} \end{cases} \qquad i\in\mathcal{N},\ p\in\mathcal{P} \tag{48}$$

$$\sum_{p\in\mathcal{P}} x_{ij}^p \le \sum_{f\in\mathcal{F}}\sum_{r\in\mathcal{R}^f} \delta_{ij}^{rf} u_r^f y_r^f \qquad (i,j)\in\mathcal{A} \tag{49}$$

$$\sum_{r\in\mathcal{R}^f} y_r^f \le n_f \qquad f\in\mathcal{F} \tag{50}$$

$$\sum_{r\in\mathcal{R}^f} \beta_i^r y_r^f = 0 \qquad i\in\mathcal{N},\ f\in\mathcal{F} \tag{51}$$

$$\sum_{r\in\mathcal{R}^f} \delta_r^h y_r^f \le a_h \qquad h\in\mathcal{H} \tag{52}$$

$$y_r^f \ge 0 \text{ et entier} \qquad r\in\mathcal{R}^f,\ f\in\mathcal{F} \tag{53}$$

$$x_{ij}^p \ge 0 \qquad (i,j)\in\mathcal{A},\ p\in\mathcal{P} \tag{54}$$

Les contraintes (48) garantissent la conservation des flux pour chaque produit, tandis que les contraintes liantes (49) limitent le flux sur chaque arc de vol à la capacité de l'avion qui emprunte ce parcours. Les contraintes (50) et (52) assurent la réalisabilité des solutions par rapport au nombre d'avions de chaque type disponibles et aux capacités d'atterrissage aux hubs, respectivement. Les contraintes (51) représentent les restrictions d'équilibre des avions : le nombre d'avions de chaque type qui atterrissent à un emplacement (parcours de livraison) doit être égal au nombre d'avions qui décollent du même lieu (parcours de ramassage).

Dans Kim *et al.* [105], les auteurs examinent les formulations fondées sur les arcs, les chemins et les arbres et choisissent la dernière option puisque celle-ci réduit considérablement la taille du problème. Pour résoudre la relaxation linéaire de la formulation qui en découle, les auteurs combinent des heuristiques afin de réduire encore davantage la taille du problème, des inégalités dans une méthode de coupes et la génération de colonnes. On utilise ensuite la technique branch-and-bound pour obtenir une solution en nombres entiers. Armacost, Barnhart et Ware [5] propose une approche méthodologique différente. Les auteurs tirent parti du fait que les parcours aériens passent au plus par deux points de passage et transforment la formulation du problème en définissant des variables qui représentent des combinaisons des parcours de service. Les nouvelles variables composées permettent de prendre en considération des parcours aériens multiples avec une seule variable. En outre, les variables de parcours composées rendent compte implicitement de la distribution des flux et produisent ainsi une formulation de conception pure pour laquelle on peut calculer des bornes plus solides et développer par conséquent des méthodes de résolution plus efficaces. Les résultats obtenus à partir des données fournies par une

grande entreprise de courrier exprès sont très encourageants et ils soulignent combien il est nécessaire de continuer à explorer les formulations de conception de réseaux afin d'ouvrir de nouvelles pistes et de mettre au point des méthodes de résolutions plus efficaces.

La réorganisation des services postaux allemands s'inscrit dans la même classe de problèmes, mais à une échelle plus étendue. Pour ramener le problème à des dimensions acceptables, Grünert et Sebastian [92] (voir aussi Grünert, Sebastian et Thärigen [93] et Buedenbender, Grünert et Sebastian [25]) le décomposent en plusieurs sous-problèmes : l'optimisation du réseau de poste aérienne de nuit, la conception du système de transport aux fins d'approvisionnement terrestre et de livraison et l'ordonnancement des opérations. Des modèles et des techniques de tournées de véhicules, ainsi qu'une formulation discrète de conception de réseaux dynamique sont proposés pour les tâches d'acheminement et d'ordonnancement. On décompose encore davantage la formulation de conception de réseau aérien en un problème de vol direct et en un problème de vols avec hubs ; les deux produisent des formulations de conception de réseau multiproduit avec coûts fixes et capacités comportant des contraintes supplémentaires. Pour optimiser ces formulations, les auteurs proposent des combinaisons d'heuristiques classiques, des métaheuristiques de recherche avec tabous et évolutives, ainsi que des méthodes de branch-and-bound.

4.3 Formulations dynamiques

Lorsque des horaires doivent être construits, une dimension temporelle doit être incorporée de façon explicite dans les formulations de conception du réseau de service. Pour ce faire, on représente habituellement les opérations du système durant un certain nombre de périodes de temps en utilisant un réseau espace-temps.

La représentation du réseau physique est répliquée à chaque période. Quittant son point d'origine durant une période donnée, un service arrive plus tard à d'autres terminaux (et en repart, dans le cas d'arrêts intermédiaires). Les services génèrent ainsi des liens de service temporels entre différents terminaux à différentes périodes de temps. Les liens temporels qui connectent deux représentations du même terminal à deux périodes de temps différentes peuvent représenter le temps requis par les activités du terminal ou l'attente des marchandises pour le prochain départ. Les coûts associés aux arcs de ce réseau sont similaires à ceux utilisés dans les formulations statiques de la sous-section précédente. On peut avoir recours à des arcs supplémentaires pour rendre compte des pénalités imposées pour une arrivée trop précoce ou trop tardive.

Ici encore, on fait appel à deux types de variables de décision. Des variables de design en nombres entiers sont associées à chaque service. Ces variables, contraintes à des valeurs de $\{0, 1\}$, indiquent si le service quitte ou non à une heure spécifiée. Quand plusieurs départs peuvent avoir lieu durant la même période de temps, on doit utiliser des variables générales (non négatives) en nombres entiers (il faut noter qu'on peut toujours utiliser uniquement des variables $\{0, 1\}$ en réduisant les périodes de temps de façon appropriée).

Des variables continues sont utilisées pour représenter la distribution des flux de marchandises à travers ce réseau de service. Les formulations obtenues sont des modèles de conception de réseaux similaires à ceux présentées dans la sous-section précédente, mais sur des réseaux nettement plus étendus en raison de la dimension temporelle.

La taille considérable du réseau espace-temps ainsi que les contraintes additionnelles généralement imposées par la dimension temporelle rendent cette classe de problèmes encore plus difficile à résoudre que les problèmes statiques. On ne trouve, en fait, qu'un nombre limité de travaux portant sur ce sujet et on y propose surtout des méthodes de résolution heuristiques pour des applications particulières (par exemple, Dall'Orto *et al.* [28], Equi *et al.* [66], Farvolden et Powell [69, 70], Haghani [95]). Un effort significatif de recherche est requis dans ce domaine.

4.4 Gestion de flottes

La plupart des tâches opérationnelles sont exécutées dans un environnement hautement dynamique où le facteur temps joue un rôle important. Notons, en particulier, que les répartiteurs doivent répondre aux demandes des clients en temps réel ou quasi réel, ce qui réduit considérablement le temps de réaction disponible pour la prise de décision. Le facteur temps est également déterminant pour la représentation des processus d'arrivée de l'information et décisionnel. Il est en effet nécessaire d'intégrer dans les décisions d'aujourd'hui les conséquences possibles de ces décisions sur l'état futur du système, ainsi que sur la productivité, le rendement et les décisions associés. La modélisation de l'aspect dynamique des opérations est rendue encore plus complexe par la stochasticité inhérente à la plupart des systèmes. Si le trafic est plus lent que prévu, les véhicules risquent d'arriver en retard chez les clients ou au terminal. Des demandes de clients prévues peuvent ne pas se matérialiser tandis que des demandes imprévues peuvent devoir être satisfaites. Les approvisionnements en véhicules vides prévus aux dépôts peuvent ainsi fluctuer et il peut se révéler nécessaire de procéder à des déplacements à vide supplémentaires.

La planification opérationnelle des activités et la gestion des ressources (équipages, personnel en général, véhicules, unités de traction, etc.) en temps (quasi) réel est d'une grande importance tant pour le transport personnalisé que pour les transporteurs avec consolidations. C'est aussi un vaste domaine que nous ne pouvons qu'effleurer. Nous nous penchons donc essentiellement sur une seule de ces questions, particulièrement pertinente et cruciale pour le transport de marchandises, soit la gestion des mouvements à vide, que nous illustrons à l'aide de deux modèles pour la gestion d'une flotte de conteneurs. Pour plus de détails concernant les modèles et méthodes pour le niveau opérationnel d'activité des entreprises de transport, on peut consulter : Cordeau, Toth et Vigo [36], Christiansen *et al.* [31], Crainic [38, 40], Crainic et Laporte [49], Crainic et Kim [48], Dejax et Crainic [59], Delorme, Roy et Rousseau [60], Powell [143, 144], Powell, Jaillet et Odoni [150], Powell et Topaloglu [154, 155], Powell, Bouzaïene-Ayari et Simaõ [145].

La nécessité pour les transporteurs de fret de déplacer des véhicules vides résulte des écarts observés à la plupart des sites entre la demande et l'offre pour chaque type de produits, écarts qui se traduisent par une accumulation de véhicules vides dans des secteurs où ils ne sont pas requis et par des pénuries de véhicules dans d'autres régions qui en ont besoin. Il faut alors soit trouver des charges supplémentaires, soit déplacer des véhicules vides afin de acheminer le « bon » nombre et type de véhicules là où ils sont requis pour satisfaire la demande connue et prévue des périodes subséquentes. Cette opération, qu'on désigne sous le nom de *repositionnement*, ou équilibrage, est une composante majeure de ce qu'on appelle la gestion de flottes. Dans sa forme la plus générale, la gestion de flottes englobe l'éventail complet des questions de planification et de gestion, depuis l'approvisionnement en unités motrices et en véhicules jusqu'à la répartition des véhicules, à l'établissement des horaires des équipes et aux travaux d'entretien. Toutefois, il arrive souvent que ce terme désigne un ensemble plus restreint d'activités : allocation des véhicules en fonction des requêtes des clients et repositionnement des véhicules vides. C'est cette définition que nous utiliserons dans ce qui suit.

Le mouvement de véhicules vides ne contribue pas directement aux bénéfices de l'entreprise, mais il est essentiel à la poursuite de ses activités. Par conséquent, il faut tenter de minimiser les mouvements à vide dans les limites imposées par les exigences en matière de demande et de service. Historiquement, les premiers modèles ont été élaborés pour le problème d'équilibrage des wagons de chemin de fer et ils utilisaient des formulations de transport : compte tenu des estimations d'offre et de demande futures en wagons vides (flotte homogène) aux gares de triage du réseau, ainsi que du coût du mouvement des wagons entre chaque paire de terminaux, optimiser la distribution des wagons vides pour minimiser le coût total (Leddon et Wrathall [117] et Misra [126], par exemple). La considération explicite de la perspective temporelle a permis d'accroître considérablement les capacités de modélisation. Un réseau temps-espace représente les divers chemins que peuvent emprunter les véhicules pour atteindre leur destination à l'heure spécifiée, y compris les décisions d'attendre aux terminaux. La formulation qui en résulte prend la forme d'un modèle déterministe de réseau de transbordement dynamique, où les flux sont optimisés de telle manière que le coût total est minimisé ou la rentabilité du système est maximisée. À la suite des contributions pionnières de White [183, 182] et de White et Bomberault [184], de nombreux modèles sur la distribution des véhicules vides ont adopté cette forme. La considération explicite des incertitudes dans les modèles de distribution de véhicules vides a constitué la prochaine contribution méthodologique d'importance. Le premier effort majeur dans ce sens a été réalisé par Jordan et Turnquist [103] pour le transport ferroviaire. La structure du modèle est cette fois encore un réseau multiproduit dynamique qui intègre de manière explicite la stochasticité de l'offre, de la demande et des temps de déplacement. La formulation d'optimisation non linéaire qui en découle est résolue au moyen de l'algorithme de Frank-Wolfe [72]. Les nombreux développements qui ont suivi, dont la plupart s'appliquaient au transport ferroviaire et routier, sont revus dans les articles mentionnés précédemment.

Crainic, Gendreau et Dejax [46] ont proposé une série de modèles pour l'allocation et la gestion d'une flotte hétérogène de conteneurs. Pour garder les dimensions de ce cha-

pitre dans des limites raisonnables, nous ne présentons ici que la formulation déterministe multiproduit. Le problème a été énoncé dans le contexte de la portion terrestre du transport international de conteneurs, tel qu'il est décrit dans la section 3.1. Le système se compose d'un certain nombre de terminaux portuaires, de dépôts intérieurs et d'installations de clients. Les navires arrivent dans les ports avec à leur bord des conteneurs chargés et vides de divers types et dimensions. Les conteneurs chargés sont livrés à leurs points de destination par rail et par camion selon les contrats respectifs, tandis que les conteneurs vides sont disponibles pour toute livraison à des clients à proximité du port ou à des fins de repositionnement. Les clients qui reçoivent des conteneurs chargés les déchargent et signalent qu'ils peuvent être ramassés et transportés à un terminal désigné (portuaire ou intérieur). De même, les clients qui ont besoin de conteneurs vides de types particuliers pour des envois futurs les reçoivent d'un terminal intérieur ou portuaire. Si le type ou le nombre de conteneurs requis par le client n'est pas disponible pour être livré dans la fenêtre de temps spécifiée par celui-ci, l'entreprise peut louer des conteneurs (ou en emprunter auprès de sociétés partenaires) ou remplacer un type de conteneur par un autre. Les conteneurs vides disponibles pour être distribués aux clients proviennent ainsi de quatre sources : les clients qui déchargent leurs marchandises, le repositionnement à l'échelle mondiale, la location et la substitution. À l'intérieur même des systèmes, deux mécanismes permettent le repositionnement des conteneurs : 1) les mouvements massifs d'équilibrage déterminés par la planification stratégique / tactique des opérations (section 3.1) et 2) les décisions relatives au dépôt de réception des conteneurs qui deviennent vides chez les clients. Le problème est dynamique en ce sens que les demandes varient dans le temps et que le transport et les opérations des clients peuvent nécessiter plus qu'une période (un jour). Il est stochastique en raison des variations des demandes des clients, de la possibilité de demandes provenant de clients « inconnus », des variations du temps requis par les clients pour décharger et retourner les conteneurs et des dommages aux conteneurs (les variations du temps de parcours n'ont pas été modélisées dans cet article). L'objectif est de minimiser le coût total d'exploitation, y compris les substitutions et les pénuries pour un horizon de planification multipériode.

La formulation est définie pour un horizon de planification discrétisé en T périodes, $t = 1, 2, ..., T$. Soit $\mathcal{H}$ l'ensemble de ports et $\mathcal{D}$ l'ensemble de dépôts terrestres. L'ensemble des requêtes des clients (demande) pour des conteneurs de type p est noté $[\mathcal{I}^p$, avec un nombre d_i^p de conteneurs de type p requis par un client $i \in \mathcal{I}^p$, à être livrés dans la fenêtre de temps $[\Delta_i$. L'ensemble de requêtes de clients qui peuvent être exécutées (c'est à dire que les conteneurs requis de type p arrivent dans les limites Δ_i) en livrant à partir du dépôt $j \in \mathcal{D} \cup \mathcal{H}$ à la période t est noté $\mathcal{I}_j^{pt}$, tandis que $\mathcal{J}_i^{pt}$ indique l'ensemble de dépôts j qui peuvent envoyer des conteneurs de type p pour qu'ils arrivent durant la période t chez le client i. $\mathcal{S}^{pt}$ représente l'ensemble de clients offreurs libérant o_s^{pt} conteneurs vides de type p durant la période t ; $\mathcal{S}_j^{pt}$ est l'ensemble des clients offreurs à partir desquels des conteneurs de type p peuvent être acheminés au terminal j dans la période t. Le nombre de conteneurs vides de type p requis au port $h \in \mathcal{H}$ durant la période t à des fins d'exportation est noté d_h^{pt}, tandis que o_h^{pt} représente le volume de conteneurs vides de type p arrivant au port h durant la période t et c_{pr} indique le nombre de conteneurs de type p requis pour remplacer un conteneur de type r.

Les coûts unitaires de transport pour un conteneur de type p sont c_{ji}^{pt} du dépôt j au client demandeur $i \in \mathcal{I}_j^{pt}$, c_{sj}^{pt} depuis un client offreur $s \in \mathcal{S}_j^{pt}$ jusqu'au dépôt j et c_{jk}^{pt} entre les dépôts $j, k \in \mathcal{D} \cup \mathcal{H}$. Le coût de substitution au dépôt j à la période t d'un conteneur de type r par un conteneur de type p est noté c_j^{prt}, tandis que le coût d'inventaire du même conteneur au dépôt j à la période t est noté c_j^{pt}. Finalement, le coût de location (ou d'emprunt) d'un conteneur de type p, au dépôt j à la période t est $\bar{c}_j^{pt}$, ce qui représente aussi le coût de pénalité au port h quand la demande d_h^{pt} ne peut pas être satisfaite. Avec les variables de décision

x_{ji}^{pt} : Nombre de conteneurs équivalents au type p alloués durant la période t depuis le dépôt j au client i ;

x_j^{pt} : Nombre de conteneurs de type p alloués en tant que type p durant la période t depuis le dépôt j ;

x_j^{prt} : Nombre de conteneurs de type p remplacés par des conteneurs de type r durant la période t au dépôt j ;

x_{sj}^{pt} : Nombre de conteneurs de type p ramassés chez le client s durant la période t et livrés au dépôt j durant la période $\bar{t} \geq t$;

w_j^{pt} : Nombre de conteneurs de type p en stock au dépôt j à la fin de la période t ;

w_{jk}^{pt} : Nombre de conteneurs de type p repositionnés depuis le dépôt j au dépôt k à la période t ;

b_j^{pt} : Nombre de conteneurs de type p loués au dépôt j durant la périodet ;

le modèle minimise le coût total du système

$$\sum_t \sum_{p \in \mathcal{P}} \sum_{j \in \mathcal{H} \cup \mathcal{D}} \left(\sum_{s \in \mathcal{S}_j^{pt}} c_{sj}^{pt} x_{sj}^{pt} + \sum_{i \in \mathcal{I}_j^{pt}} c_{ji}^{pt} x_{ji}^{pt} + \sum_{k \in \mathcal{H} \cup \mathcal{D}} c_{jk}^{pt} w_{jk}^{pt} + \sum_{r \in \mathcal{P}} c_j^{prt} x_j^{prt} + c_j^{pt} w_j^{pt} + \bar{c}_j^{pt} b_j^{pt} \right) \tag{55}$$

sujet à des contraintes de non-négativité sur toutes les variables, des bornes sur les flux d'équilibrage et

$$\sum_{t \in \Delta_i} \sum_{j \in \mathcal{J}_i^{pt}} x_{ji}^{pt} = d_i^p \quad i \in \mathcal{I}^p \tag{56}$$

$$\sum_{\bar{t} \geq t} \sum_{j \in \mathcal{S}_{jp\bar{t}}} x_{sj}^{p\bar{t}} = o_s^{pt} \quad s \in \mathcal{S}^{pt},\ p \in \mathcal{P},\ t = 1, 2, ..., T \tag{57}$$

$$x_j^{pt} + \sum_{r \in \mathcal{P}} c_{pr} x_j^{prt} - \sum_{i \in \mathcal{I}_j^{pt}} x_{ji}^{pt} = 0 \quad j \in \mathcal{D} \cup \mathcal{H},\ p \in \mathcal{P},\ t = 1, 2, ..., T \tag{58}$$

$$w_j^{pt} + x_j^{pt} + x_j^{prt} + \sum_{k \in \mathcal{H} \cup \mathcal{D}} w_{jk}^{pt} - w_j^{pt-1} - \sum_{s \in \mathcal{S}_j^{pt}} x_{sj}^{pt} - \sum_{\bar{t} \leq t} \sum_{k \in \mathcal{H} \cup \mathcal{D}} w_{kj}^{p\bar{t}} - b_j^{pt} = 0$$
$$j \in \mathcal{D},\ p \in \mathcal{P},\ t = 1, 2, ..., T \tag{59}$$

$$w_h^{pt} + x_h^{pt} + x_h^{prt} + \sum_{k \in \mathcal{H} \cup \mathcal{D}} w_{hk}^{pt} + d_h^{pt} - w_{hk}^{pt} - \sum_{s \in \mathcal{S}_h^{pt}} x_{sh}^{pt} - \sum_{\bar{t} \leq t} \sum_{k \in \mathcal{H} \cup \mathcal{D}} w_{hk}^{p\bar{t}} - \sigma_h^{pt} - b_h^{pt} = 0$$

$$h \in \mathcal{H},\ p \in \mathcal{P},\ t = 1, 2, ..., T \quad (60)$$

où les contraintes (56) et (57) spécifient les conditions de l'offre et de la demande aux clients, les équations (4.4) et (4.4) sont les contraintes de conservation des flux aux dépôts terrestres et aux ports, respectivement, tandis que les équations (58) tiennent les comptes des « vrais » conteneurs et les conteneurs de substitution alloués aux dépôts.

Le modèle d'allocation dynamique de conteneurs vides de Cheung et Chen [29] envisage le problème du point de vue d'une société de navires porte-conteneurs qui assure des services réguliers entre un nombre donné de ports. Les navires transportent des conteneurs chargés qui génèrent des bénéfices et, si l'espace le permet, des conteneurs vides à repositionner afin de pouvoir satisfaire une future demande prévue. Il est assumé qu'un seul navire voyage entre deux ports à chaque période et que la demande en conteneurs chargés entre les ports ne dépasse pas la capacité du navire. Un seul produit – les conteneurs vides – est pris en considération. L'horaire des navires qui relient les ports dans $\mathcal{H}$ est supposé fixe pour l'horizon de planification $t = 1, 2, ..., T$. Comme dans les formulations de Crainic, Gendreau et Dejax, on suppose que le modèle est utilisé dans un mode d'horizon fuyant : le modèle donne une solution (qui tient compte de l'impact des décisions actuelles sur l'état futur du système), la solution suggérée pour la première période est exécuté et ensuite, à la période suivante, le modèle est utilisé de nouveau avec l'information mise à jour.

Pour un conteneur, le coût du transport depuis le port $i \in \mathcal{H}$ au port $j \in \mathcal{H}$, quittant à la période t est noté c_{ij}^t, tandis qu'à un port i à la période t, les coûts de chargement / déchargement (du navire) sont l_i^t et u_i^t, le coût d'inventaire est $[c_i^t$ et le coût de location (ou d'emprunt) est $\bar{c}_i^t$. Le temps de transport du port i au port j est noté τ_{ij} et le revenu généré par la satisfaction d'une unité de demande au port i durant la période t est donné par R_i^t.

Trois sources de stochasticité sont prises en compte dans le modèle : $[\kappa_{ij}^t$, la capacité résiduelle pour des conteneurs vides à bord du navire voyageant du port i au port j, quittant à la période t, δ_i^t, la demande pour des conteneurs au port i durant la période t et σ_i^t, la quantité de conteneurs disponibles au port i à la période t (avant le déchargement des navires). Les variables de décision (non négatives) sont :

$x_i^{l,t}$: Nombre de conteneurs chargés au port i durant la période t pour toute destination ;

$x_i^{u,t}$: Nombre de conteneurs déchargés au port i durant la période t depuis tout autre port ;

$x_{ij}^{c,t}$: Nombre de conteneurs à repositionner à partir du port i au port j quittant durant la période t ;

$x_{ij}^{p,t}$: Nombre de conteneurs actuellement à bord du navire au port i, à la période t, avec comme destination le port j ;

$x_i^{h,t}$: Nombre de conteneurs entreposés au port i à la période t ;

$x_i^{r,t}$: Nombre de conteneurs loués, utilisés au port i à la période t pour satisfaire la demande ;

$x_i^{d,t}$: Nombre de conteneurs à soi, utilisés au port i à la période t pour satisfaire la demande.

L'objectif, stochastique, est de minimiser le coût total prévu tout en maximisant le revenu généré par l'exécution de la demande

$$\text{Minimiser } E\left[\sum_t \sum_{i\in\mathcal{H}} \left(l_i^t x_i^{l,t} + u_i^t x_i^{u,t} + c_i^t x_i^{h,t} + \bar{c}_i^t x_i^{r,t} + \sum_{j\in\mathcal{H}} c_{ij}^t x_{ij}^{c,t} - R_i^t \delta_i^t \right)\right], \tag{61}$$

alors que les relations suivantes définissent la dynamique du système et les contraintes du modèle :

$$\sum_{k\in\mathcal{H}} x_{ki}^{c,t-\tau_{ij}} - x_i^{u,t} - \sum_{j\in\mathcal{H}} x_{ij}^{p,t} = 0 \qquad i\in\mathcal{H},\ t=1,2,...,T \tag{62}$$

$$x_i^{h,t} + \delta_i^t - x_i^{h,t-1} - \sigma_i^t - x_i^{h,t} = 0 \qquad i\in\mathcal{H},\ t=1,2,...,T \tag{63}$$

$$\sum_{j\in\mathcal{H}} x_{ij}^{c,t} - x_i^{l,t} - \sum_{j\in\mathcal{H}} x_{ij}^{p,t} = 0 \qquad i\in\mathcal{H},\ t=1,2,...,T \tag{64}$$

$$x_{ij}^{c,t} \le \kappa_{ij}^t \qquad i\in\mathcal{H},\ t=1,2,...,T \tag{65}$$

$$x_i^{r,t} = \delta_i^t - x_i^{d,t} \qquad i\in\mathcal{H},\ t=1,2,...,T. \tag{66}$$

Les relations (65) spécifient que la demande doit être satisfaite. À chaque port et à chaque période, les équations (62) calculent le volume de conteneurs déchargés des navires (la différence entre les volumes totaux qui arrivent avec les navires et en partent), les relations (63) représentent les équations d'inventaire qui respectent aussi l'exigence selon laquelle la demande doit être satisfaite et les équations (64) calculent le nombre de conteneurs qui sont repositionnés. Les contraintes (65) respectent la capacité résiduelle pour les mouvements de repositionnement et les équations (66) définissent le nombre de conteneurs loués en tant que différence entre le nombre de conteneurs à soi disponibles et la demande. Le modèle peut être vu comme un problème de réseaux avec capacités aléatoires sur les arcs et est présenté sous la forme d'un programme stochastique à deux étapes. Deux méthodes de résolution (quasi-gradients et approximation hybride) sont proposées et comparées au moyen d'une expérimentation à horizon fuyant. Rappelons, pour conclure cette section, l'approche méthodologie initialement connue sous le nom de « Logistic Queuing Networks » (Powell *et al.* [149], Powell et Carvalho [146, 147, 148]). Basée sur des principes de programmation dynamique et de décomposition, le raffinement de cette méthodologie a permis d'obtenir des résultats tout à fait impressionnants pour des applications réalistes de gestion de ressources en contexte dynamique et stochastique (Godfrey and Powell [89, 90], Powell and Topaloglu [155, 154], Powell [144], Powell, Bouzaïene-Ayari, and Simaõ [145]).

5 Problèmes de tournées de véhicules

Un des problèmes centraux n distribution de marchandises et, de façon générale, en logistique est le problème de tournées de véhicules (PTV), ou *vehicle routing problem* selon la terminologie anglaise. Ce problème consiste à déterminer un ensemble optimal de circuits de distribution ou de ramassage à partir d'un ou plusieurs dépôts afin de desservir un ensemble de clients ou de fournisseurs. Les différentes possibilités existant tant au niveau du réseau routier, de la flotte de véhicules employés, que des contraintes sur la constitution des tournées, induisent autant de variantes de ce problème et en font sa richesse. Récemment, le livre de Toth et Vigo [179] a permis de faire un survol complet des recherches menées sur le PTV jusqu'à la fin du vingtième siècle.

Dans ce chapitre nous focaliserons trois variantes principales du PTV : le problème de tournées de véhicules avec contraintes de capacité (PTVC), le problème de tournées de véhicules avec contraintes de capacité et de longueur (PTVCL) et le problème de tournées de véhicules avec fenêtres de temps (PTVFT). Nous mettrons en exergue les modèles et méthodes qui nous apparaissent comme les plus pertinentes actuellement pour ces différents problèmes.

5.1 Définitions et complexité

Le PTV est défini sur un graphe orienté ou non $G = (\mathcal{V}, \mathcal{A})$ ou $G = (\mathcal{V}, \mathcal{E})$ où $\mathcal{V} = \{v_0, \ldots, v_n\}$ est un ensemble de sommets représentant les dépôts ou les clients et $\mathcal{A} = \{(v_i, v_j) \text{ tel que } i \neq j,\ v_i, v_j \in \mathcal{V}\}$ ou $\mathcal{E} = \{\{v_i, v_j\} \text{ tel que } i \neq j,\ v_i, v_j \in \mathcal{V}\}$ un ensemble d'arcs ou d'arêtes les reliant. A chaque arc ou arête est associé un coût c_{ij}. Sauf indication contraire, nous supposerons dans la suite que le graphe G est complet et non orienté.

La version standard du PTV est le problème de tournées de véhicules avec contraintes de capacité (PTVC) qui a été décrit pour la première fois par Dantzig et Ramser en 1959 [55]. Dans le PTVC, on associe à chaque client $v_i \in \mathcal{V}\backslash\{v_0\}$ un volume q_i de marchandises. Le PTVC consiste à déterminer un ensemble optimal de tournées issues d'un unique dépôt représenté par le noeud v_0 et desservant un ensemble de client tel que la distance totale parcourue par les véhicules soit minimisée en vérifiant que pour chaque véhicule le volume total des biens transportés ne dépasse pas sa capacité Q.

Une extension classique du PTVC est le problème de tournées de véhicules avec contraintes de capacité et de longueur (PTVCL). On définit alors pour chaque client $v_i \in \mathcal{V}\backslash\{v_0\}$ un temps de service s_i et pour chaque arête un temps de parcours t_{ij}. Par rapport au PTVC, le PTVCL vise à prendre en compte la restriction additionnelle imposant que la longueur de chaque tournée n'excède pas une valeur spécifiée L. Une seconde variante fréquemment considérée du PTVC est le problème de tournées de véhicules avec fenêtres

de temps (PTVFT). Dans le PTVFT, on associe à chaque sommet $v_i \in \mathcal{V}$ une fenêtre de temps $[a_i, b_i]$ et on impose de plus par rapport au PTVC que le client $v_i \in \mathcal{V}$ soit visité au cours de cet intervalle.

Le PTVC, le PTVCL ainsi que le PTVFT sont des problèmes NP-durs. La complexité du PTVC a été étudiée par Lenstra et Rinnooy Kan [118]. Les complexités du PTVCL et du PTVFT en découlent. Il est à noter que lorsque la taille de la flotte est fixée, les problèmes d'existence d'une solution admissible sont des problèmes NP-complets.

5.2 Extensions classiques

Nous n'avons décrit ci-dessus que les trois variantes les plus communes du PTV. Or ce problème en possède de très nombreuses issues, notamment, des multiples applications pratiques existantes. Parmi les diverses versions du PTV, nous pouvons distinguer sans prétendre être exhaustif les cas suivants :

- au niveau du réseau de distribution : le réseau peut être asymétrique ou symétrique ce qui équivaut à définir le problème sur un graphe orienté ou non ; la flotte peut être répartie entre plusieurs dépôts ;
- au niveau de la clientèle : la livraison et la collecte de biens peuvent être réalisées soit successivement soit simultanément ; l'accès chez un client peut être limité à un sous-ensemble des véhicules ;
- au niveau de la flotte de véhicules : le nombre m de véhicules utilisés peut être fixé ; la flotte peut être hétérogène, chaque véhicule possédant ses propres caractéristiques ;
- au niveau des chauffeurs : les contraintes sociales telles que la gestion des temps de pauses et de conduite peuvent être considérées ; chaque chauffeur peut ne pouvoir conduire que certains véhicules ;

De même, l'objectif considéré peut différer. Si la minimisation de la distance totale parcourue est le plus souvent considérée, on rencontre également les cas suivants :

- minimisation de la taille de la flotte ;
- minimisation des coûts d'utilisation, un coût fixe étant alors associé à l'utilisation de chacun des véhicules ;
- minimisation de la violation des contraintes, en particulier pour les contraintes de fenêtres de temps ;
- maximisation des gains, un gain étant lié à la visite d'un client lorsque visiter la totalité de la clientèle est impossible.

Finalement, plus récemment, un autre type de généralisation a consisté à considérer des versions multi-objectifs notamment dans des cas pratiques. Pour un survol de la littérature sur ce sujet, nous renvoyons le lecteur à [104].

6 Modèles et méthodes exactes pour le PTV

Dans cette section, nous allons commencer par présenter trois modélisations mathématiques des problèmes abordés pour lesquelles on suppose que la taille de la flotte est fixée à m. Le premier et le troisième modèles sont suffisamment généraux pour pouvoir être étendus pour traiter la plupart des variantes. Par contre, le second modèle est spécifique au PTVC dans le cas où le réseau est symétrique. Il faut souligner que les deux dernières modélisations ont conduit à des méthodes qui font partie des plus efficaces pour résoudre exactement le PTVC et le PTVFT. Nous en donnerons les caractéristiques principales dans la seconde partie de cette section. Ne prétendant pas couvrir toutes les types de modèles et les méthodes existants, nous renvoyons le lecteur intéressé au livre de Toth et Vigo [179] ainsi qu'à l'article de Laporte et Nobert [113].

6.1 Modèle de flot à trois indices

Ce premier modèle, qui correspond à ce qui peut sembler une modélisation « naturelle »du problème, est défini sur un graphe orienté $G = (\mathcal{V}, \mathcal{A})$. Il consiste à définir une variable binaire x_{ijk} égale à 1 si le véhicule k parcourt l'arc $(i, j) \in A$. Le modèle pour le PTVC s'écrit :

$$\text{Minimiser} \quad \sum_{k=1}^{m} \sum_{(v_i, v_j) \in \mathcal{A}} c_{ij} x_{ijk} \tag{67}$$

$$\text{sujet à} \quad \sum_{k=1}^{m} \sum_{v_j \in \mathcal{V}} x_{ijk} = 1 \qquad v_i \in \mathcal{V} \setminus \{v_0\}, \tag{68}$$

$$\sum_{v_i \in \mathcal{V}} x_{i\ell k} = \sum_{v_j \in \mathcal{V}} x_{\ell jk} \qquad v_\ell \in \mathcal{V} \setminus \{v_0\},\ 1 \le k \le m \tag{69}$$

$$\sum_{v_j \in \mathcal{V} \setminus \{v_0\}} x_{0jk} = 1 \qquad 1 \le k \le m \tag{70}$$

$$\sum_{v_i \in \mathcal{V} \setminus \{v_0\}} x_{i0k} = 1 \qquad 1 \le k \le m \tag{71}$$

$$\sum_{v_i \in \mathcal{V} \setminus \{v_0\}} \sum_{v_j \in \mathcal{V}} q_i x_{ijk} \le Q \qquad 1 \le k \le m \tag{72}$$

$$\sum_{v_i \in \mathcal{S}, v_j \notin \mathcal{S}} x_{ijk} \ge \sum_{v_j \in \mathcal{V}} x_{\ell jk} \qquad \mathcal{S} \subset \mathcal{V} \setminus \{v_0\}, \ell \in \mathcal{S},\ 1 \le k \le m \tag{73}$$

$$x_{ijk} \in \{0, 1\} \qquad (v_i, v_j) \in \mathcal{A},\ 1 \le k \le m \tag{74}$$

Dans cette formulation, les contraintes (68) imposent que chaque client soit desservi une et une seule fois. Les contraintes (69) sont les contraintes de conservation de flot

alors que les contraintes (70) et (71) assurent que chaque tournée démarre et s'achève au dépôt. Notons que si la taille de la flotte est uniquement bornée par m, il suffit de remplacer dans ces contraintes les égalités par des inégalités. Les contraintes (72) sont les contraintes de capacité et les contraintes (73) les contraintes d'élimination de sous-tours. Finalement les contraintes (74) sont les contraintes de binarité.

Ce type de formulation a souvent été utilisé comme modèle mathématique de nouvelles variantes du PTV. En effet, ayant un indice par véhicule, on peut aisément spécifier l'ensemble des paramètres du problème par rapport à chacun d'eux et il est en général facile d'exprimer de nouvelles contraintes portant sur les tournées grâce aux variables utilisées. On obtient par exemple directement une formulation du PTVCL en ajoutant les contraintes suivantes sur la longueur des tournées.

$$\sum_{v_i \in \mathcal{V} \setminus \{v_0\}} \sum_{v_j \in \mathcal{V}} s_i x_{ijk} + \sum_{(v_i,v_j) \in \mathcal{A}} t_{ij} x_{ijk} \leq L \qquad 1 \leq k \leq m \tag{75}$$

De même, une formulation du PTVFT en découle si on introduit de nouvelles variables u_{ik} représentant l'heure d'arrivée du véhicule k chez le client v_i. On peut alors remplacer les contraintes d'élimination de sous-tours (73) par les contraintes suivantes :

$$a_i \sum_{v_j \in \mathcal{V}} x_{ijk} \leq u_{ik} \leq b_i \sum_{v_j \in \mathcal{V}} x_{ijk} \qquad v_i \in \mathcal{V},\ 1 \leq k \leq m \tag{76}$$

$$u_{ik} + s_i + t_{ij} - u_{jk} \leq T(1 - x_{ijk}) \qquad (v_i, v_j) \in \mathcal{A},\ 1 \leq k \leq m \tag{77}$$

$$u_{ik} \geq 0 \qquad v_i \in \mathcal{V},\ 1 \leq k \leq m \tag{78}$$

où T est une constante suffisamment grande.

6.2 Modèle de flot pour le PTVC symétrique

Ce second modèle a été proposé pour le PTV symétrique par Laporte *et al.* [114]. De façon à obtenir une écriture plus simple du modèle, on indice les données pertinentes par les arêtes. Ainsi, le PTV étant défini sur un graphe $G = (\mathcal{V}, \mathcal{E})$, le coût de l'arête $e \in \mathcal{E}$ est noté c_e. On définit ensuite $\delta(S)$ comme étant l'ensemble des arêtes adjacentes à l'ensemble de sommets S. En particulier $\delta(\{v_i\})$ ou $\delta(v_i)$ est l'ensemble des arêtes adjacentes au noeud v_i. En associant à chaque arête $e \in \mathcal{E}$ une variable binaire x_e égale à 1 si un véhicule l'emprunte, le modèle pour le PTVC s'écrit alors :

$$\text{Minimiser} \quad \sum_{e \in \mathcal{E}} c_e x_e \tag{79}$$

$$\text{sujet à} \quad \sum_{e \in \delta(v_i)} x_e = 2 \qquad v_i \in \mathcal{V} \setminus \{v_0\}, \tag{80}$$

$$\sum_{e \in \delta(v_0)} x_e = 2m \tag{81}$$

$$\sum_{e \in \delta(\mathcal{S})} x_e \geq 2\, r(\mathcal{S}) \qquad \mathcal{S} \subset \mathcal{V} \setminus \{v_0\},\ 2 \leq |\mathcal{S}| \leq n-2 \tag{82}$$

$$x_e \in \{0,1,2\} \qquad e \in \delta(v_0) \tag{83}$$

$$x_e \in \{0,1\} \qquad e \in \mathcal{E} \setminus \delta(v_0) \tag{84}$$

Dans cette formulation, les contraintes (80) et (81) sont les contraintes de degré alors que les contraintes (82) sont les contraintes de capacité. Dans leur expression, $r(\mathcal{S})$ peut être interprété comme étant le nombre de véhicules distincts nécessaires à la desserte de l'ensemble S de sommets (voir section 6.4.1). Finalement, les contraintes (83) modélisent le fait qu'une arête issue du sommet associé au dépôt peut être parcourue deux fois dans le cas d'une tournée n'incluant qu'un sommet, une fois ou pas du tout alors que les contraintes (84) sont les contraintes de binarité associées aux autres variables. Comme nous le verrons dans la section 6.4, c'est à partir de ce modèle qu'ont été proposées des méthodes faisant partie des plus performantes pour résoudre le PTVC symétrique.

6.3 Modèle de partitionnement d'ensemble

Dans cette formulation, introduite par Balinski et Quandt [15], on définit T comme étant l'ensemble de toutes les tournées admissibles. Puis, à toute tournée $t \in \mathcal{T}$, on associe son coût c_t ainsi que pour tout sommet v_i, un coefficient binaire a_{it} égal à 1 si le sommet v_i appartient à la tournée t. En considérant des variables binaires x_t qui prennent la valeur 1 si la tournée $t \in \mathcal{T}$ est choisie dans la solution, nous obtenons la formulation :

$$\text{Minimiser} \quad \sum_{t \in \mathcal{T}} c_t x_t \tag{85}$$

$$\text{sujet à} \sum_{t \in \mathcal{T}} a_{it} x_t = 1 \qquad v_i \in \mathcal{V} \setminus \{v_0\}, \tag{86}$$

$$\sum_{t \in \mathcal{T}} x_t = m \tag{87}$$

$$x_t \in \{0,1\} \qquad t \in \mathcal{T} \tag{88}$$

Dans ce modèles les contraintes (86) spécifient que chaque sommet est visité une et une seule fois alors que la contrainte (87) impose une taille de flotte de m. Cette dernière

contrainte peut être simplement omise s'il n'y a pas de restriction sur le nombre de véhicules à utiliser. Finalement, les contraintes (88) sont les contraintes de binarité.

Ce modèle est très général puisque toutes les contraintes sur la constitution des tournées n'apparaissent pas explicitement. En fait, hormis celle de résoudre le problème de partitionnement d'ensemble, la difficulté principale avec ce type de modèle est la génération de l'ensemble des tournées T. En effet, pour un sous-ensemble de clients donné, cela revient typiquement à résoudre un problème de voyageur de commerce avec des contraintes additionnelles dépendant de la variante du PTV considérée. Dès lors, suivant que l'on détermine un sous-ensemble de tournées a priori ou que l'on génère les tournées au cours de la résolution, cela donne naissance à des heuristiques (voir section 7.1) ou à des méthodes exactes parmi les plus performantes notamment pour le PTVFT (voir section 6.5).

6.4 Méthodes de séparations et coupes pour le PTVC

Basés sur la modélisation présentée à la section (6.2), les algorithmes de séparations et coupes (*Branch & Cut*) constituent une des approches les plus performantes pour résoudre le PTVC symétrique de façon exacte. De manière générale, un algorithme de séparations et coupes consiste, à partir d'une modélisation en nombres entiers du problème, à considérer une relaxation du modèle dans laquelle les restrictions sur l'intégrité sur les variables ainsi qu'éventuellement certaines contraintes sont relâchées. Etant donné la solution optimale de la relaxation, deux cas sont considérés : soit la solution est une solution admissible du problème traité et dans ce cas l'algorithme s'achève, soit il s'agit d'une solution non admissible. A l'aide de procédures, appelées *procédures de séparation*, on identifie alors des contraintes valides du problème qui sont violées par la solution courante. Ces contraintes sont alors ajoutées au modèle relaxé et le processus est réitéré. Si aucune coupe ne peut être générée ou si la borne induite par la résolution du modèle relaxé ne varie plus de façon significative, un processus de séparation-évaluation est déclenché. A chaque noeud de l'arbre la borne est calculée selon le procédé expliqué précédemment. Dans la prochaine sous-section, nous présenterons q uelques unes des contraintes valides qui ont été proposées. Puis, nous décrirons les procédures de séparation les plus courantes. Cette section est basée sur les travaux de Augerat [8], de Ralphs *et al.* [157], Achuthan, Caccetta et Hill [1] et finalement de Lysgaard, Lechford et Eglese [121]. Le livre de Toth et Vigo [179] consacre un chapitre détaillé à ce sujet.

6.4.1 Inégalités valides

Les contraintes que nous présentons dans cette section sont des contraintes valides pour le polytope du PTVC défini comme l'enveloppe convexe de toutes les solutions du PVTC impliquant m tournées. La taille du polytope étant une fonction complexe des

paramètres du problème, il est difficile de prouver si ces contraintes valides sont des facettes.

Le premier groupe est constitué par les contraintes proposées pour le PVC symétrique. En effet, Naddef et Rinaldi [128] ont montré que toute contrainte valide pour le polytope du PVC symétrique, mise sous forme triangulaire serrée, est une contrainte valide pour le PTVC. Pour $a, x \in \Re^{|\mathcal{E}|}$ et $b \in \Re$, une inégalité $ax \geq b$ est dite sous forme triangulaire serrée si pour tout triplet d'arêtes $((v_i, v_j), (v_i, v_k), (v_j, v_k))$ l'inégalité triangulaire est vérifiée : $a_{(i,j)} \leq a_{(i,k)} + a_{(k,j)}$ et pour tout sommet $v_i \in \mathcal{V}$, il existe deux sommets $v_{i_1}, v_{i_2} \in \mathcal{V}$ tels que : $a_{(i_1,i_2)} = a_{(i,i_1)} + a_{(i,i_2)}$. Naddef et Rinaldi [128] ont montré que toute contrainte valide pour le PVC symétrique peut se mettre sous forme triangulaire serrée.

Le second groupe de contraintes est celui des contraintes de capacité. Les contraintes de capacité sont les restrictions (82) du modèle présenté à la section (6.2). Dépendant de la valeur donnée à $r(S)$, cet ensemble de contraintes prend des noms différents. Nous verrons dans la sous-section suivante que ces différentes classes se distinguent au niveau des procédures de séparation mise en oeuvre. Si on considère la valeur la plus faible pour le membre de droite : $r(\mathcal{S}) = \sum_{v_i \in \mathcal{S}} q_i/Q$, on parle de contraintes de capacité *fractionnaires*. Si on prend l'arrondi supérieur de la valeur précédente : $r(\mathcal{S}) = \lceil \sum_{v_i \in \mathcal{S}} q_i/Q \rceil$, on parle de contraintes de capacité *arrondie*. Etant donné S, une valeur valide pour $r(S)$ est la valeur optimale du problème de mise en boîtes (*bin packing*) pour lequel les poids des objets sont égaux à q_i pour $v_i \in S$ et la capacité des boîtes égale à Q. On parle alors de contraintes de capacité *faible*. Finalement, si on donne à $r(\mathcal{S})$ la valeur du nombre de véhicules nécessaires à la desserte des sommets de $\mathcal{S}$ lorsque l'on dispose d'une flotte de m véhicules, ces contraintes sont les contraintes de capacité *globales*.

Une généralisation des contraintes de capacité sont les contraintes de boîte (*framed capacity inequaliy*) introduites par Augerat [8]. Cette famille de contraintes est définie pour tout $\mathcal{H}, \mathcal{S}_1, \ldots, \mathcal{S}_t \subseteq \mathcal{V} \setminus \{v_0\}$ tels que :

$$\mathcal{S}_i \subset \mathcal{H} \qquad i = 1, \ldots, t \tag{89}$$

$$\mathcal{S}_i \cap \mathcal{S}_j = \emptyset \qquad i \neq j \; i, j = 1, \ldots, t \tag{90}$$

$$\sum_{v_k \in S_i} q_k \leq Q \qquad i = 1, \ldots, t \tag{91}$$

$$\tag{92}$$

La contrainte de boîte s'exprime alors de la façon suivante :

$$\sum_{e \in \delta(\mathcal{H})} x_e + \sum_{i=1}^{t} \sum_{e \in \delta(\mathcal{S}_i)} x_e \geq 2(t + r(\mathcal{H}|\mathcal{S}_1, \ldots, \mathcal{S}_t)) \tag{93}$$

où $r(\mathcal{H}|\mathcal{S}_1, \ldots, \mathcal{S}_t)$ est la valeur optimale du problème de mise en boîtes où les boîtes ont une capacité Q et où l'on associe à chaque sous-ensemble S_i un objet de poids égal à la demande totale des sommets de $\mathcal{S}_i$, et à chaque sommet de $\mathcal{H} \setminus (\bigcup_{i=1}^{t} \mathcal{S}_i)$ un objet de poids égal à la demande en ce sommet. Intuitivement, cette contrainte signifie que si les

contraintes de capacité sont serrées pour chaque $\mathcal{S}_i \subset \mathcal{H}$ alors $r(\mathcal{H}|\mathcal{S}_1, \ldots, \mathcal{S}_t)$ constitue une borne inférieure sur le nombre de véhicules nécessaire à la desserte de $\mathcal{H}$. Augerat [8] décrit d'autres formes plus complexes de contraintes de ce type.

Plusieurs autres familles de contraintes valides comme les contraintes de capacité généralisées, les contraintes d'étoile ou encore les contraintes d'hypotour ont été proposées. Récemment, Achuthan, Caccetta et Hill [1] ont suggéré de nouvelles familles de contrainte qui éliminent des solutions admissibles mais pas de solutions optimales. Un exemple de contraintes de ce type est le suivant. Soit $\mathcal{S} \subseteq \mathcal{V} \setminus \{v_0\}$ tel que $\sum_{v_i \in \mathcal{S}} q_i \leq Q$, avec $|\mathcal{S}| \geq 2$:

$$\sum_{e=(v_i,v_j)\ v_i,v_j \in \mathcal{S}} x_e + \sum_{e=(v_0,v_j)\ v_j \in \mathcal{S}} x_e \leq |\mathcal{S}| + 1 \tag{94}$$

6.4.2 Méthodes de séparation

Les procédures de séparation pour les coupes venant du PVC sont analogues à celles proposées dans ce cadre. Naddef et Thienel [129] [130] ont publié récemment une description détaillée de celles-ci. Dans cette section nous considérerons plus particulièrement les méthodes de séparation qui ont été proposées pour les contraintes de capacité. Les contraintes de capacité fractionnaires sont simples à séparer. En effet, les identifier revient à résoudre un problème de flot. Par contre, le problème de séparation de toutes les autres formes de contraintes de capacité présentées à la section précédente est un problème NP-complet [8]. Dès lors, différentes heuristiques ont été suggérées.

Partant du graphe support, c'est à dire du sous-graphe de G pour lequel $x_e > 0$, l'heuristique la plus simple consiste à identifier les composantes connexes du graphe support privé du sommet v_0 puis à vérifier si la contrainte de capacité est violée pour chacune des composantes. Si ce n'est pas le cas, on considère l'ensemble $\mathcal{S}$ des sommets d'une des composantes d'où l'on exclut le sommet pouvant conduire à une violation et on procède à une nouvelle vérification. On répète si nécessaire le procédé tant qu'il demeure des sommets dans $\mathcal{S}$. Au besoin, on considère les sommets d'une autre composante.

Une seconde approche est une procédure gloutonne. Elle consiste, à partir d'un ensemble $\mathcal{S}$ de sommets consistant initialement par exemple en un sommet de $\mathcal{V} \setminus \{v_0\}$, à inclure dans $\mathcal{S}$ le sommet $v \in \mathcal{V} \setminus (\{v_0\} \cup \mathcal{S})$ pour lequel $\sum_{e=(v,v_j)\ v_j \in \mathcal{S}} x_e$ est maximum, puis à vérifier si la contrainte de capacité définie pour $\mathcal{S}$ est violée. A partir de cette méthode, Augerat *et al.* [7] ont proposé une adaptation de la méta-heuristique tabou pour identifier des contraintes de capacité violées.

Ces heuristiques ainsi que des extensions du même type sont reprises dans les approches d'Augerat [8], de Ralphs *et al.* [157], et de Lysgaard, Lechford et Eglese [121]. Ralphs *et al.* [157] ont proposé une autre approche basée sur un algorithme de décomposition. Le but de cette méthode est de déterminer l'ensemble $\mathcal{T}$ de tournées tel que la solution

fractionnaire courante s'écrive comme une combinaison convexe de ces tournées. Si une telle décomposition est obtenue, alors une contrainte de capacité est nécessairement violée pour une des tournées de $\mathcal{T}$ sinon une contrainte de séparation peut être générée d'après le théorème de Farkas.

Les expériences numériques sont effectuées sur un ensemble d'instances communes [101]. Les comparaisons sont toutefois délicates. Augerat *et al.* [7] ont implanté un algorithme impliquant plusieurs classes de contraintes valides dont les contraintes de capacité, de capacité généralisées, de peigne et d'hypotour. Ils ont obtenu les solutions optimales d'instances comportant jusqu'à 135 sommets et 7 véhicules en des temps atteignant au maximum 20570 secondes sur une sun sparc 20. L'implémentation en parallèle de leur algorithme n'incluant que des routines de séparation pour les contraintes de capacité a conduit Ralphs *et al.* [157] à résoudre des instances comportant jusqu'à 100 sommets et 8 véhicules en des temps de calcul pouvant atteindre environ 2 millions de secondes. La méthode de séparation et coupe de Lysgaard, Lechford et Eglese [121] inclut comme celle d'Augerat *et al.* plusieurs classes de coupe dont certaines plus fortes que celles considérées précédemment [7]. Ces chercheurs obtiennent les solutions optimales pour des problèmes incluant jusqu'à 120 sommets en des temps de calculs allant jusqu'à 12000 secondes. Comme ils le soulignent, leur approche semble être plus efficace sur les instances les plus dures alors que l'algorithme de Ralphs *et al.* est plus rapide sur les problèmes plus faciles. Finalement, Achuthan, Caccetta et Hill [1] n'ont repris qu'un sous-ensemble des instances tests classiques et ont calculé la matrice des distances différemment ce qui rend la comparaison difficile.

6.5 Méthodes de génération de colonnes pour le PTVFT

L'une des approches les plus efficaces pour résoudre le PTVFT de façon exacte est d'avoir recours à un algorithme de type génération de colonnes. Cette approche a notamment été mise en oeuvre par Desrochers, Desrosiers et Solomon [62]. Elle repose sur la formulation du PTVFT sous forme de partitionnement d'ensemble (voir section 6.3) et procède de la façon suivante. On définit le problème maître comme étant le problème de partitionnement d'ensemble défini pour un sous-ensemble $\mathcal{T}'$ de tournées de $\mathcal{T}$. A une itération donnée, la relaxation linéaire du problème maître est résolue, puis étant donné les valeurs des variables duales, une tournée, c'est à dire une colonne de coût réduit négatif, est recherchée. L'identification d'une telle colonne se fait au travers de la résolution du sous-problème qui est, dans ce cas, un problème de plus court chemin élémentaire avec contraintes de capacité et de fenêtre de temps. Ce dernier est un cas particulier du problème de plus court chemin élémentaire avec contraintes de ressource qui est un problème NP-dur [65]. Une relaxation classique consiste à en considérer la version non élémentaire. Etant donné que les coûts des arcs peuvent être négatifs suite à leur modification par les variables duales, la solution du problème du plus court chemin non élémentaire sous contraintes de ressource peut comporter des cycles. Toutefois, les contraintes de capacité et de fenêtre de temps font que la longueur du chemin est bornée.

Desrochers [61] par exemple a proposé un algorithme basé sur la programmation dynamique. Une amélioration de tels algorithmes consiste à éviter les 2-cyles, c'est à dire les cycles de la forme (v_i, v_j, v_i) [98].

Récemment, Kohl *et al.* [109] ont proposé une amélioration de cette approche consistant à ajouter au problème maître des contraintes violées par la solution fractionnaire courante. Les restrictions qu'ils génèrent sont des contraintes de k-chemins autrement dit des contraintes de capacité *globales* (voir section 6.4.1). La difficulté d'identifier de telles contraintes venant notamment de la présence des contraintes de fenêtres de temps les a conduit à se restreindre au cas où $k = 2$.

Si la résolution de la relaxation linéaire du problème maître ne débouche pas sur l'obtention d'une solution entière, un processus de séparation-évaluation est déclenché. Toutefois, à chaque noeud de l'arbre, le problème maître est résolu selon le procédé expliqué précédemment. Une des difficultés qui se pose lors de la définition d'une séparation est de déterminer sur quelles variables le branchement va être effectué. Ici, brancher sur les variables du problème maître n'est pas évident. En effet, si imposer une tournée dans une solution est simple, tenir compte du fait qu'une tournée spécifique n'est pas présente n'est pas évident au niveau de la résolution du sous-problème. Dès lors, Desrochers, Desrosiers et Solomon [62] ont proposé d'effectuer la séparation sur les arcs du graphe G. On impose alors d'une part qu'un arc donné soit présent dans la solution et d'autre part qu'il soit absent. Ces deux restrictions peuvent facilement être inclues dans la résolution du sous-problème.

L'évaluation des algorithmes pour le PTVFT est réalisée sur les 87 instances tests de Solomon [100]. Celles-ci se répartissent en 3 classes de problèmes (R1, C1, RC1). Chaque classe contient entre 8 et 12 instances de 25, 50 et 100 clients. Pour les classes R, les clients sont répartis uniformément alors que pour les classes C ils sont répartis en groupes. Les problèmes de la classes RC constituent une catégorie mixte par rapport à la répartition. Les problèmes de catégorie 1 ont des fenêtres de temps serrées et la capacité des véhicules est de 200 ce qui conduit à utiliser de 3 à 20 véhicules. L'approche de Kohl *et al.* [109] est sans doute l'une des meilleures. Sur les 87 instances, ces chercheurs ont en effet résolus 70 problèmes à l'optimum en des temps de calcul d'au plus 1216 secondes sur une machine à 100 MHz.

7 Méthodes heuristiques pour le PTV

Comme nous l'avons vu précédemment Le PTVC, le PTVCL ainsi que le PTVFT sont des problèmes NP-durs et, à ce jour, les chercheurs ont proposé des approches exactes qui ne permettent de résoudre que des instances de taille modeste (voir section 6). Lorsque l'on considère des cas plus complexes faisant intervenir des contraintes multiples ou que l'on doit résoudre des problèmes de plus grandes tailles, des approches heuristiques sont

nécessaires. Depuis les années soixante, de nombreuses heuristiques dites *classiques*, c'est à dire non basées sur les méta-heuristiques, ont été proposées par les chercheurs. A cet ensemble de méthodes se sont adjointes depuis la fin des années quatre-vingts des approches basées sur les méta-heuristiques.

Dans cette section nous présentons un survol des approches heuristiques proposées pour le PTV. Aujourd'hui, les heuristiques classiques ne sont plus compétitives ni en termes de qualité de solution ni en termes de temps de calcul. En effet, comme nous le verrons dans la section 7.2 les adaptations les plus récentes des méta-heuristiques l'emportent sur ces deux critères. Il n'en demeure pas moins que les heuristiques classiques restent d'actualité à deux points de vue. Tout d'abord les plus rapides d'entre elles sont utilisées dans des méthodes basées sur les méta-heuristiques soit pour fournir une ou plusieurs solutions initiales soit pour reconstruire une solution admissible à partir d'une solution partielle. Ensuite certains des principes à partir desquelles elles avaient été développées sont aujourd'hui repris au sein des meilleures approches basées sur les méta-heuristiques. Le survol que nous présentons est partiel et résulte de choix motivés par la pertinence des méthodes et le caractère récent de certaines d'entre elles. Nous renvoyons le lecteur aux livres de Toth et Vigo [179] et de Pirlot et Teghem [177] pour une revue plus complète d'approches plus anciennes.

7.1 Heuristiques classiques

Les heuristiques classiques peuvent être réparties en trois classes :
- les méthodes constructives ;
- les algorithmes à deux phases ;
- les procédures d'améliorations

Nous allons dans cette section décrire un sous-ensemble de méthodes dans chacune de ses classes. Notre choix est essentiellement fondé sur l'utilisation qui en est faite aujourd'hui dans les approches basées sur les méta-heuristiques.

7.1.1 Méthodes constructives

Le plus connu de ces algorithmes est sans doute la méthode des économies (*savings*) de Clarke et Wright [33]. Partant d'une solution initiale où chaque client est desservi sur une tournée, elle consiste à calculer pour toute paire de sommets $(v_i, v_j) \in \mathcal{E}, v_i, v_j \in \mathcal{V} \setminus \{v_0\}$ l'économie e_{ij} réalisée en allant directement de v_i à v_j plutôt qu'en passant par le dépôt :

$$e_{ij} = c_{i0} + c_{0j} - c_{ij} \tag{95}$$

Les économies étant classées de manière décroissante, on procède dans la version parallèle à chaque itération à l'union des deux tournées correspondant à la plus grande économie et préservant l'admissibilité de la solution. Dans la version séquentielle, on considère une

tournée spécifique que l'on essaye d'étendre. Cette heuristique a été proposée originellement pour le PTVC et le PTVCL. Elle fut par la suite étendue aux différentes variantes qui ont été considérées du PTV. Solomon en a proposée une version pour le PTVFT [172].

La méthode d'insertion proposée par Mole et Jameson [127] fait intervenir deux critères lors de la construction des tournées. Le premier est un critère d'insertion d'un sommet $v_k \in \mathcal{V} \setminus \{v_0\}$ entre deux clients consécutifs $v_i, v_j \in \mathcal{V}$ sur une tournée :

$$\alpha(v_i, v_k, v_j) = c_{ik} + c_{kj} - \lambda c_{ij} \tag{96}$$

Dépendant de la valeur donnée au paramètre $\lambda, \alpha(v_i, v_k, v_j)$ s'interprète différemment. Dans le cas où $\lambda = 1$, il correspond au surcoût généré par le détour par le sommet v_k par rapport à une liaison directe. Le second critère est donné par :

$$\beta(v_i, v_k, v_j) = \mu c_{0k} - \alpha(v_i, v_k, v_j) \tag{97}$$

Si on fixe la valeur du paramètre μ à 2, $\beta(v_i, v_k, v_j)$ s'interprète comme le gain engendré par l'insertion du client v_k entre v_i et v_j par rapport à une desserte directe depuis le dépôt. L'heuristique de Mole et Jameson construit les tournées de manière séquentielle. Partant de la tournée constituée par le sommet non desservi le plus éloigné du dépôt, elle consiste à calculer pour chaque sommet non inclus sur une tournée, le coût associé à sa meilleure insertion en utilisant le premier critère, puis à déterminer le meilleur sommet à insérer en utilisant le second critère. Suite à une insertion, l'heuristique 3-Opt [119] est appliquée pour améliorer le coût de la tournée courante. Proposée pour le PTVC et le PTVCL, Solomon [172] a étendu ce type d'approche au PTVFT. Il a notamment redéfinit le critère $\alpha(v_i, v_k, v_j)$ de façon à prendre en compte le décalage temporel crée par l'insertion du sommet v_k :

$$\alpha(v_i, v_k, v_j) = \alpha_1(c_{ik} + c_{kj} - \lambda c_{ij}) + \alpha_2(t_j(v_k) - t_j)) \tag{98}$$

avec $\alpha_1 + \alpha_2 = 1, \alpha_1, \alpha_2 \geq 0$ et $t_j(v_k), t_j$ étant respectivement l'heure d'arrivée du véhicule au sommet v_j après et avant l'insertion du sommet v_k.

Comme cela a été dit en début de section, aucune de ces méthodes n'est compétitive en terme de qualité de solution sur les instances tests de la littérature. Ainsi, si on considère les résultats obtenus sur un ensemble de 14 problèmes tests pour le PTVC et le PTVCL comprenant de 51 à 200 sommets proposés par Christofides, Mingozzi et Toth ([32], [99]) pour lesquels seules deux solutions optimales sont connues [94], [91]. La meilleure implantation de la méthode des économies combinée avec l'heuristique 3-opt [119] fournit des solutions à 6,72% des solutions optimales ou des meilleures solutions connues [179]. Pour la méthode de Mole et Jameson, l'implantation récente de Prins [156] donne un écart moyen de 10,76%. Les temps de calcul sont pour ces deux approches très réduits. Ainsi la méthode des économies fournit une solution en 0,2 seconde sur une station sun Ultrasparc 10 de 42 Mflops.

L'évaluation des heuristiques dans ce chapitre pour le PTVFT est considérée par rapport aux 56 instances tests de Solomon [100] comportant 100 clients. Outre les 29

problèmes de la catégorie 1 décrits ci-dessus, on inclut 27 problèmes de la catégorie 2 répartis dans les classes (R2, C2, RC2) pour lesquels la capacité des véhicules est portée à 1000 pour R2 et RC2, et à 700 pour C2. Dans son article, Solomon [172] précise qu'en matière d'heuristiques, l'objectif est tout d'abord de minimiser le nombre de tournées puis leur coût total. L'adaptation de la méthode des économies par Solomon [172] n'est pas probante et la meilleure méthode d'insertion, qu'il a proposée, génère 454 tournées au total sur les 56 instances. Cela doit être mis en comparaison avec les 405 tournées obtenues par la meilleure méthode actuelle [19].

7.1.2 Méthodes à deux phases

Les méthodes à deux phases se répartissent en deux classes d'algorithmes suivant l'ordre dans lequel les deux phases sont effectuées : les algorithmes de partition puis de parcours et les algorithme de parcours d'abord puis de partition. Dans la première catégorie se trouve des méthodes très connues telles que l'algorithme de balayage et les algorithmes de pétales.

L'algorithme de balayage (*sweep*) dû à Gillett et Miller [84] repose sur l'hypothèse qu'à chaque sommet du graphe sont associées des coordonnées polaires définies en prenant le dépôt pour origine et un sommet choisi de façon quelconque. Les sommets sont ensuite classés par coordonnée angulaire croissante. Partant du sommet de plus petite coordonnée angulaire non affecté, les sommets sont inclus successivement dans le groupe courant tant que les contraintes sont vérifiées. Lorsqu'une des contraintes est violée, la constitution du groupe est stoppée et l'on passe à l'élaboration du groupe suivant. Le processus est répété jusqu'à ce que tous les sommets soient inclus dans un groupe. Les groupes de sommets étant formés, un PVC est résolu pour chacun d'eux. Si la satisfaction des contraintes de capacité peut se faire aisément lors de l'inclusion d'un sommet dans un groupe, celles portant sur la longueur de la tournée ou sur les fenêtres de temps nécessitent la résolution d'un PVC lors de chaque insertion [172]. Fisher et Jaikumar [71] ont proposé une méthode similaire qui consiste à résoudre un problème d'affectation généralisée au cours de la première phase puis un PVC pour chaque véhicule. Afin de déterminer le coût d'affectation d'un sommet à un véhicule, Fisher et Jaikumar associe à chaque véhicule un sommet germe, le coût d'affectation devenant le coût d'insertion entre le dépôt et ce sommet fictif.

Les algorithmes de pétales sont issus d'une modélisation du PTV sous la forme d'un problème de partitionnement d'ensemble (voir section 6.3). Ces routines consistent à générer un sous-ensemble de $\mathcal{T}$ de tournées (appelées pétales) a priori intéressantes puis à résoudre le modèle de partition d'ensemble restreint à ces tournées. La meilleure des méthodes de ce type pour le PTVC et le PTVCL est la méthode 2-pétales proposée par Renaud, Boctor et Laporte [159]. Elle consiste à générer non seulement des tournées individuelles mais également des blocs de deux tournées qui soit s'entrecroisant soit sont inclues l'une dans l'autre.

Parmi les algorithme de parcours d'abord puis de partition, l'approche proposée par Beasley [17] pour le PTVC et le PTVCL consiste à résoudre dans un premier temps un PVC sur $\mathcal{V} \setminus \{v_0\}$ puis à déterminer une partition optimale de ce tour. Cette dernière est obtenue en résolvant un problème de plus court chemin dans un graphe sans cycle.

Si on considère le PTVC et le PTVCL, la méthode de balayage implantée par Renaud, Boctor et Laporte [159] produit sur les 14 problèmes tests mentionnés ci-dessus des solutions dont la valeur est en moyenne à 7,09 % des valeurs des solutions optimales ou des meilleures solutions connues. Ce pourcentage est ramené à 2,38 % pour les solutions produites par l'algorithme 2-pétales, ce qui représente l'un des meilleurs pourcentages pour les approches classiques. Pour ces deux méthodes les temps de calculs moyens sont de 1,76 et 3,48 secondes respectivement sur une station sun Ultrasparc 2 de 4,2 Mflops. Pour le PTVFT, l'adaptation de la méthode de balayage par Solomon [172] fournit 476 tournées au total sur les 56 instances tests.

7.1.3 Méthodes d'amélioration

Deux classes de méthodes d'amélioration pour le PTV peuvent être distinguées. La première comprend les méthodes qui essayent d'améliorer individuellement chacune des tournées. Il s'agit d'algorithmes exacts ou heuristiques pour le PVC ou le problème du voyageur de commerce avec fenêtre de temps (PVCFT). Plusieurs heuristiques classiques pour le PVC telles que les algorithmes de type k-opt proposés par Lin [119], la méthode Or-Opt [137], ou encore la procédure GENIUS [74] ont été adaptés au PVCFT [165], [173], [76].

La seconde inclut des heuristiques d'amélioration de la solution courante en modifiant plusieurs tournées simultanément. Elles sont typiquement basées sur la définition d'un mécanisme d'échange d'arêtes ou d'arcs entre tournées. Pour le PTVC et le PTVCL, Van Breedam [181] a considéré trois types de modifications, que l'on retrouve dans des approches basées sur les méta-heuristiques : le premier consiste à déplacer au plus k clients consécutifs d'une tournée vers une autre ; le second vise à échanger k clients consécutifs d'une tournée avec ℓ clients consécutifs d'une seconde tournée ; le troisième combine ces deux opérations en examinant laquelle engendre la meilleure amélioration.

Pour le PTVFT, un type de modification s'apparentant à la seconde modification a été proposé pour le PTVFT par Thompson et Psaraftis [178]. Russell [166] propose une approche constructive parallèle partant du nombre minimal de tournées. Son approche inclut une procédure d'amélioration qui vise à échanger des clients entre deux tournées. Pour garder un voisinage de taille raisonnable, il considère pour chaque sommet uniquement les deux meilleures tournées possibles. Cordone et Wolfler-Calvo [37] ont proposé une heuristique du même type qui consiste à effectuer une recherche locale à partir de différentes solutions initiales fournies par la méthode d'insertion de Solomon [172]. Un algorithme de type k-opt avec k=2, 3 est appliquée suivie d'une procédure de type 5-

échanges qui vise à supprimer une tournée en réinsérant ses sommets dans les autres. Pour que la recherche ne reste pas bloquée dans un optimum local, une modification de la fonction objectif est utilisée. Elle consiste à prendre pour objectif la durée totale des tournées en omettant les temps d'attente.

Pour le PTVC et le PTVCL, Van Breedam a arrondi les distances ce qui rend la comparaison impossible avec les meilleures solutions publiées pour lesquelles les distances sont des réels. On peut toutefois indiquer que la meilleure stratégie est la troisième stratégie. Les résultats se situent à 4,18 % en moyenne des meilleures solutions obtenues au cours de différentes expériences numériques. Les temps de calcul correspondant varient entre 3 et 421 secondes sur un ordinateur équipé d'un processeur 80386. Sur les 56 instances du PTVFT, Thompson et Psaraftis [178] obtiennent 439 tournées, Russell [166] 424 et Cordone et Wolfler-Calvo [37] 422. Comme nous le verrons à la section 7.2.2, l'écart par rapport à de nombreux algorithmes basés sur les méta-heuristiques reste important.

7.2 Méta-heuristiques

Depuis les années quatre-vingts dix, les heuristiques les plus performantes pour le PTV sont basées sur diverses méta-heuristiques. De très nombreuses adaptations ont été proposées ([179], [177]). Dans cette section, nous ne reprenons que celles qui nous apparaissent aujourd'hui comme étant les plus performantes. Nous distinguerons les adaptations pour le PTVC et le PTVCL de celles développées pour le PTVFT. Toutefois, pour chaque problème, nous considérerons leurs différents composants en mettant en évidence les points communs et les relations par rapport aux heuristiques classiques.

7.2.1 Méta-heuristiques pour le PTVC et le PTVCL

Méthodes basées sur la recherche tabou Nous nous restreignons aux travaux de Taillard [174], de Gendreau, Hertz et Laporte [75], de Rochat et Taillard [161], de Rego et Roucairol [158], de Tarantilis et Kiranoudis [176] et de Toth et Vigo [180].

Espace des solutions : Dans toutes les approches considérées, l'espace des solutions S est étendu. En effet, S contient non seulement les solutions admissibles pour le problème initial mais également des solutions non admissibles. Ainsi, Taillard [174] relâche les contraintes d'affectations des clients aux tournées, alors que Gendreau, Hertz et Laporte [75] relaxent les contraintes de capacité et de longueur. Dans le premier cas, un terme de pénalité fixe pour chaque sommet non visité est introduit dans la fonction objectif alors que dans le second la fonction objectif est modifiée en introduisant deux termes associés aux violations des contraintes de capacité et de longueur. Ces termes sont pondérés par des coefficients qui varient dynamiquement au cours de la recherche en fonction des violations générées. Un tel principe est repris dans les autres approches

considérées ici. Pour diminuer le temps de calcul requis, Toth et Vigo [180] ont proposé de réduire l'espace des solutions étendu décrit ci-dessus en excluant des solutions jugées a priori peu prometteuses. Après avoir observé que les arêtes inclues dans les meilleures solutions connues avait un coût nettement inférieur au coût moyen des arêtes dans le graphe, ils ont suggéré de travailler sur un graphe réduit $G' = (\mathcal{V}, \mathcal{E}')$ où $\mathcal{E}'$ est l'ensemble des arêtes adjacentes au dépôt et de celles dont le coût est inférieur à une valeur de seuil θ. La valeur de θ est obtenue en calculant le produit d'un paramètre β par le coût moyen des arêtes dans une solution générée par la méthode des économies (voir section 7.1.1). Dès lors, seules les solutions définies sur G' sont examinées. Notons qu'en faisant varier la valeur de β, une phase d'intensification ou de diversification de la recherche peut être déclenchée.

Exploration du voisinage : Taillard [174], Tarantilis et Kiranoudis [176], et Toth et Vigo [180] utilisent tous le voisinage obtenu par la stratégie de λ-échanges proposée par Osman [138]. Celle-ci consiste à sélectionner une paire de routes p et q, puis dans chacune d'elles un sous-ensemble de sommets, respectivement $\mathcal{S}_p$ et $\mathcal{S}_q$ avec $|\mathcal{S}_p| \leq \lambda$ et $|\mathcal{S}_q| \leq \lambda$, et finalement à remplacer dans p les sommets de $\mathcal{S}_p$ par ceux de $\mathcal{S}_q$ et dans q ceux de $\mathcal{S}_q$ par ceux de $\mathcal{S}_p$. $\mathcal{S}_p$ ou $\mathcal{S}_q$ pouvant être vide, ce type d'échange décrit aussi bien le transfert de sommets d'une tournée vers une autre, que l'échange de sommets entre deux tournées. La taille du voisinage croissant rapidement, la valeur de λ est fixée à 1 dans les différentes approches. De plus, Taillard [174] ne considère que les échanges de sommets avec $|\mathcal{S}_p| = |\mathcal{S}_q| = 1$. Gendreau, Hertz et Laporte [75] utilise une structure de voisinage basée sur l'utilisation de la procédure GENI proposée initialement pour le PVC [74]. Etant donné une solution courante, un mouvement consiste à déplacer un sommet v d'une tournée dans une autre soit vide soit contenant l'un des p plus proches voisins de v. L'insertion de v est une insertion généralisée dans le sens qu'elle est effectuée entre deux sommets non nécessairement consécutifs sur la tournée. Une réoptimisation locale de la tournée doit donc être effectuée. Rego et Roucairol [158] ont considéré un voisinage généré à l'aide de chaînes d'éjection. Une telle chaîne correspond à une suite de mouvements visant à déplacer un sommet d'une tournée vers une autre, chaque sommet prenant la place d'un autre qui est « 'ejecté ». Pour un client donné, le choix du sommet dont il prend la place se fait en considérant ses h plus proches voisins. Ainsi, ℓ éjections sont réalisées, permettant de passer de la solution courante à une solution voisine. Le processus se termine en insérant le dernier client éjecté entre deux autres qui minimisent le coût d'insertion simple.

Liste tabou : Dans toutes les approches, la liste tabou est définie en interdisant le mouvement inverse du mouvement réalisé pendant un certain nombre d'itérations. Ce nombre est choisi aléatoirement dans un intervalle prédéfini.

Stratégies d'intensification et de diversification : Taillard [174] a proposé une méthode d'intensification basée sur une décomposition de la solution courante en des groupes de tournées proches les unes des autres. Plus précisément, les coordonnées polaires du barycentre de chacune des tournées sont calculées. Puis, en parcourant les tournées selon la coordonnée angulaire croissante, la solution est par-

titionnée en groupes contenant approximativement le même nombre de tournées. Une décomposition supplémentaire peut être réalisée en divisant chaque groupe de tournées en régions concentriques. L'algorithme recherche tabou est ensuite appliqué à chacun des groupes.

Taillard [174] et Gendreau, Hertz et Laporte [75] ont eu recours à la même stratégie de diversification. Celle-ci consiste à déterminer pour chaque mouvement la fréquence avec laquelle il est effectué, puis à pénaliser son coût par un facteur qui dépend notamment de cette fréquence et de la taille du voisinage. Rochat et Taillard [161] ont proposé une autre stratégie appelée *mémoire adaptative* qui permet de réaliser des phases d'intensification et de diversification de la recherche. Cette approche s'apparente à la construction d'une solution du PTV par la résolution d'un problème de partitionnement d'ensemble. La mémoire adaptative est un ensemble $\mathcal{T}'$ de tournées issues de bonnes solutions du sein desquelles des tournées sont choisies afin de construire une nouvelle solution de bonne qualité. L'ensemble $\mathcal{T}'$ est tout d'abord généré en utilisant l'algorithme de Taillard [174]. Puis, une nouvelle solution est construite de façon itérative en choisissant aléatoirement des tournées dans $\mathcal{T}'$ selon une distribution de probabilité qui dépend de la qualité des solutions auxquelles elles appartiennent. Seules les tournées qui comprennent des clients pas encore desservis peuvent être retenues. Si à la fin de ce processus de construction, la solution résultante ne comprend pas certains clients, une solution admissible est construite en créant de nouvelles tournées. Cette solution est ensuite améliorée via la procédure de recherche tabou, l'ensemble $\mathcal{T}'$ étant mis à jour dynamiquement au cours de la recherche. Tarantilis et Kiranoudis [176] ont proposé une variante de cette approche qui consiste à définir $\mathcal{T}'$ comme un ensemble de bonnes séquences de sommets (*bone*) provenant de solutions de qualité. La nouvelle solution est alors construite en incluant ces séquences de sommets. Celles-ci sont sélectionnées sur base du nombre de sommets qui les constituent et de leurs fréquences d'apparition dans les meilleures solutions visitées.

Méthodes basées sur les algorithmes génétiques Jusqu'à très récemment aucune approche compétitive basée sur les algorithmes génétiques n'avait été développée pour le PTVC et le PTVCL. Les adaptations de Baker et Ayechew [13] d'une part, et de Prins [156] d'autre part comblent ce vide.

Codage d'une solution - Espace des solutions : Baker et Ayechew [13] ont proposé une adaptation des algorithmes génétiques au PTV selon une approche de type partition d'abord parcours ensuite (voir section 7.1.2). Ils utilisent en effet un codage des solutions de type affectation. Celui-ci indique pour chaque sommet le véhicule auquel il est affecté, chaque véhicule desservant à peu près constamment la même région. La population initiale est constituée de solutions obtenues à l'aide de l'heuristique de balayage et d'une variante de la méthode de Fisher et Jaikumar [71] dans laquelle le problème d'affectation généralisée est résolu heuristiquement . Au cours de la recherche, des solutions non admissibles peuvent être inclues dans la population courante. Prins [156] a, quant à lui, utilisée une approche de type parcours d'abord partition ensuite. Cela le conduit donc à utiliser un codage sous forme de cycle, chaque solution étant représentée

par un tour incluant tous les sommets. Comme dans l'approche proposée par Beasley [17], la décomposition optimale en tournées est obtenue par la résolution d'un problème de plus court chemin dans un graphe acyclique. La population initiale est obtenue est considérant les solutions générées par trois heuristiques classiques, à savoir : la méthode des économie, l'algorithme de Mole et Jameson et la méthode de balayage. Au cours de la recherche, seules des cycles correspondant à des solutions admissibles du PTV sont inclus dans la population.

Opérateurs : Baker et Ayechew utilisent des opérateurs classiques à savoir un opérateur de croisement 2-points et un opérateur de mutation basé sur l'échange de deux clients entre deux tournées. A intervalle régulier, des méthodes de recherche locale sont appliquées sur toutes les solutions de la population. Ces méthodes sont la 2-opt [119] et une recherche locale basée sur une stratégie de 1-échange [138]. Chaque solution étant un tour, Prins [156] a choisi un opérateur de croisement classique pour le PVC : l'opérateur OX. Au lieu d'appliquer un opérateur de mutation, une recherche locale basée sur une stratégie de 2-échanges [138] est réalisée suivant une certaine probabilité.

Sur les 14 problèmes tests mentionnés ci-dessus, toutes ces approches sont compétitives si on considère le critère portant sur l'écart par rapport aux solutions optimales ou aux meilleures solutions connues. En effet, alors que l'algorithme de 2-pétales de Renaud, Boctor et Laporte [159] fournissait des solutions à 2,38% en moyenne, l'adaptation de la recherche tabou de Taillard [174] obtient des solutions à 0,05% en moyenne, la méthode de Gendreau, Hertz et Laporte [75] à 0,86 % pour un paramétrage unique et à 0,20 % si on en considère plusieurs, l'algorithme de Rego et Roucairol [158] à 0,55 % , l'adaptation de Tarantilis et Kiranoudis [176] à 0,23 % et de Toth et Vigo [180] à 0,61 %. Les deux algorithmes génétiques décrits conduisent également à des solutions de qualité, puisque l'algorithme de Baker et Ayechew [13] génère des solutions dont les valeurs sont à 0,50 % en moyenne de celles des solutions optimales ou des meilleures solutions connues alors que pour la méthode proposée par Prins [156], ce pourcentage se réduit à 0,08 %. Même si les expériences numériques ont été conduites sur des ordinateurs différents, le comportement est plus contrasté si on s'intéresse au temps de calcul requis pour obtenir ces solutions. Nous complétons ici l'évaluation initiée par Prins qui consiste à comparer les différentes méthodes lorsque le nombre de MFlops des machines est connu en prenant comme référence l'algorithme de Gendreau, Hertz et Laporte [75] avec un paramétrage standard. Cette méthode étant à 1, l'algorithme de Tarantilis et Kiranoudis [176] est à 0,58, celui de Toth et Vigo [180] à 0,2, celui de Baker et Ayechew [13] à 9,13 alors que celui de Prins [156] est à 1,46. La même évaluation est malheureusement impossible pour les autres approches. Même si ces chiffres sont à prendre avec précaution, nous constatons que globalement sur les deux critères, trois approches ressortent : celles de Tarantilis et Kiranoudis [176], de Toth et Vigo [180] et de Prins [156]. Notons toutefois qu'en termes de nombre de meilleures solutions trouvées les algorithmes de Tarantilis et Kiranoudis [176], et de Prins [156] avec respectivement 9 et 10 meilleures valeurs identifiées devancent nettement l'approche de Toth et Vigo [180] qui n'en obtient que 4.

7.2.2 Méta-heuristiques pour le PTVFT

Méthodes basées sur la recherche tabou Comme précédemment, nous nous limitons à une sélection de travaux, à savoir ceux de Rochat et Taillard [161], de Taillard *et al.* [175], de Chiang et Russell [30] et de Cordeau, Laporte et Mercier [35]. Pour un survol complet, nous renvoyons le lecteur à l'article de Bräysy et Gendreau [23] sur les approches basées sur les méta-heuristiques pour le PTVFT.

Espace des solutions : Comme pour le PTVC et le PTVCL, l'espace des solutions S est étendu pour pouvoir considérer des solutions qui sont admissibles ou qui ne le sont pas. Dans leur approche, Taillard *et al.* [175] ne relaxent que les contraintes de fenêtres de temps. Dans les trois autres adaptations, les contraintes de capacité, de longueur et de fenêtres de temps sont relâchées.

Exploration du voisinage : Rochat et Taillard [161] ainsi que Cordeau, Laporte et Mercier [35] considèrent le voisinage obtenue par le déplacement d'un sommet d'une tournée à une autre en réalisant une insertion simple. Chiang et Russell [30] utilisent le voisinage obtenu par la stratégie de λ-échanges [138]. Finalement, Taillard *et al.* [175] ont recours à un mouvement plus complexe appelé *cross-exchange* qui consiste à échanger deux chaînes de sommets consécutifs entre deux tournées.

Liste tabou : Comme précédemment la liste tabou est définie en interdisant le mouvement inverse du mouvement réalisé pendant un certain nombre d'itérations. Ce nombre est fixe dans l'adaptation de Cordeau, Laporte et Mercier [35], choisi aléatoirement dans un intervalle prédéfini dans celles de Rochat et Taillard [161], et de Taillard *et al.* [175] ou varie au cours de la recherche dans l'algorithme de Chiang et Russell [30]. Plus précisément, Chiang et Russell augmentent la longueur de la liste quand des solutions identiques sont examinées au cours de la recherche ou réduisent sa longueur lorsqu'aucune solution admissible n'est identifiée.

Stratégies d'intensification et de diversification : Les stratégies d'intensification et de diversification décrites ci-dessus dans le cadre du PTVC et du PTVCL ont été reprises par les auteurs. Ainsi, Taillard *et al.* [175] appliquent une stratégie d'intensification basée sur la décomposition d'une solution en groupes de tournées sur lesquels l'algorithme tabou est appliqué. Taillard *et al.* [175] et Cordeau, Laporte et Mercier [35] utilisent la fréquence de réalisation des mouvements pour engendrer une diversification de la recherche, alors que Rochat et Taillard [161] ont recours à l'approche de la mémoire adaptative.

Post-optimisation : Plusieurs auteurs font appel une fois la recherche tabou achevée à une méthode de post-optimisation pour améliorer la qualité de la meilleure solution obtenue. Ainsi, Taillard *et al.* [175] et Cordeau, Laporte et Mercier [35] appliquent une heuristique pour le problème du PVCFT à chacune des tournée. Rochat et Taillard [161] capitalisent sur les tournées de la mémoire adaptative pour obtenir une nouvelle solution

en résolvant un problème de partitionnement d'ensemble.

Méthodes basées sur les algorithmes évolutionnaires Si de nombreuse méthodes basées sur les algorithmes génétiques ont été proposées pour le PTVFT, les travaux les plus récents ont porté sur le développement d'algorithmes évolutionnaires. Dans ces algorithmes, il n'existe pas, nous le rappelons, de codage de solutions mais une population de solution, sur laquelle sont appliquées les opérateurs génétiques, est maintenue au cours de la recherche. Nous nous concentrons dans cette section sur certaines des approches les plus récentes conduisant à des solutions de très bonne qualité. Ces méthodes ont été suggérées par Homberger et Gehring [97], Gehring et Homberger [73], Berger, Barkaoui et Bräysy [19], et finalement Mester [123]. Comme précédemment, nous renvoyons, pour plus de détail, au survol proposé par Bräysy et Gendreau [23].

Espace des solutions : Homberger et Gehring [97] et Gehring et Homberger [73] construisent une population initiale constituée de solutions admissibles en utilisant une version modifiée de l'algorithme des économies (voir section 7.1.1). Cette variante consiste à choisir aléatoirement la paire de sommets conduisant à la fusion de deux tournées. Berger, Barkaoui et Bräysy [19] considèrent dans leur approche deux populations qui évoluent concurremment au cours de la recherche. Ces populations contiennent des solutions admissibles ou non. Leur évolution est conduite selon différents objectifs. La première évolue de façon à minimiser la distance totale alors que la seconde évolue de façon à minimiser le coût total de violations des contraintes. Les populations initiales sont construites en utilisant la méthode d'insertion de Liu et Shen [120]. Mester [123] réduit la population à une solution admissible unique qui est générée à partir de la solution précédente par mutation uniquement. La solution initiale est construite à l'aide d'une heuristique basée sur la méthodes des économies (voir section 7.1.1) et sur l'heuristique de Russell [166]. A intervalles réguliers au cours de l'élaboration de la solution, une procédure d'amélioration de 1-échange [138] est appelée.

Opérateurs : Homberger et Gehring [97] utilisent les opérateurs classiques : Or-opt [137], 2-opt* [141], 1-échange [138] comme opérateurs visant à réduire la longueur totale des tournées. Ils ont recours également à un opérateur consistant à déplacer au plus k clients consécutifs d'une tournée vers une autre afin de vider une tournée et donc de réduire le nombre de véhicules. Cet opérateur est appliqué selon une des caractéristiques du codage de la solution dans [97] par contre il est utilisé systématiquement dans [73]. Berger, Barkaoui et Bräysy [19] utilisent comme opérateurs de croisement des opérateurs basés sur une des méthodes d'insertion de Solomon [172] et sur la méthode d'insertion de Liu et Shen [120]. Le premier opérateur construit une nouvelle solution à partir d'un certain nombre de tournées d'une solution et de sommets proches de celles-ci dans l'autre solution. Le second opérateur considère une solution avec un nombre de tournées fixé. Les sommets, pour lesquels les contraintes d'heure de visite sont violées, sont retirés ainsi qu'un ensemble de sommets additionnel. Ces sommets sont ensuite réinsérés dans la solution. Leur méthode inclut cinq opérateurs de mutation. Quatre de ces opérateurs sont relativement classiques. Deux opérateurs procèdent à des échanges de sommets entre

tournées selon le principe du λ-échange [138]. Un opérateur vise à améliorer la solution en considérant les sommets pour lesquels la contrainte de fenêtre de temps n'est pas respectée alors que l'autre considère une paire de tournées proches l'une de l'autre. Un opérateur vise à réordonnancer les clients sur une tournée tandis que l'autre essaye d'éliminer la tournée la plus courte en réinsérant les sommets visités sur les autres. Ces deux derniers opérateurs utilisent les méthodes d'insertion de Solomon [172] et de Liu et Shen [120]. Le cinquième opérateur est plus original. Il s'agit d'une recherche à grand voisinage [168] qui consiste à retirer de la solution un ensemble de sommets incluant ceux pour lesquels la contrainte liée à la fenêtre de temps est violée. Ces sommets sont ensuite insérés dans la solution suivant une stratégie de type séparation-évaluation en contrôlant la taille de l'arbre. Mester [123] utilise également une recherche à grand voisinage qui consiste dans un premier temps à retirer un ensemble de sommets des tournées puis à les réinsérer à l'aide de la procédure d'insertion utilisée pour la construction de la solution initiale. Les sommets sont retirées selon trois stratégies : aléatoirement, par rapport à la distance au sommet v_0, ou encore en choisissant un sommet dans chaque tournée.

HYBRIDATION : Gehring et Homberger [73] combinent leur algorithme évolutionnaire [97] avec une méthode de recherche tabou. Cette approche consiste à utiliser comme solution initiale de la recherche tabou la solution produite par l'algorithme évolutionnaire [97]. A chaque itération, le voisinage est constitué de solutions générées à l'aide des opérateurs mentionnés ci-dessus.

Autres approches Parmi les autres approches qui ont été proposées pour les PTVFT nous décrirons ici les méthodes suggérées récemment par Bräysy, Hasle et Dullaert [24] et par Le Bouthillier et Crainic [115, 116]. Bräysy, Hasle et Dullaert [24] ont proposé une méthode de recherche locale consistant dans un premier temps à construire une population de solutions comportant un petit nombre de véhicules, puis à appliquer une méthode de recherche locale basée sur l'opérateur de croissement *cross-exchange* [175]. Finalement, la meilleure solution obtenue est améliorée à l'aide d'une méthode de recherche locale à seuil qui consiste à accepter de se déplacer vers une solution voisine si la variation de la valeur de la fonction objectif est inférieure à une valeur de seuil. Cette valeur est abaissée au cours de la recherche. Les solutions voisines de la solution courante sont obtenues en appliquant soit l'opérateur Genicross, soit l'opérateur IOPT. L'opérateur Genicross combine la méthode d'insertion Geni [74] et l'opérateur *cross-exchange* [175]. Il consiste à considérer tous les échanges possibles de chaînes de sommets consécutifs entre deux tournées. Comme dans GENI, l'insertion d'un segment au sein d'une tournée peut se faire entre deux sommets non consécutifs. IOPT est un opérateur de type *cross-exchange* qui agit sur une seule tournée à la différence de ce dernier. L'inversion du sens de parcours de la chaîne de sommets déplacée est également considérée.

Le Bouthillier et Crainic [115] ont proposé une approche coopérative selon une stratégie de *mémoire centrale* [41]. Dans une telle stratégie, plusieurs méta-heuristiques explorent de façon indépendante l'espace des solutions et échangent les bonnes solutions identifiées pendant la recherche (et leur contexte, éventuellement). Les communications se

font, de façon asynchrone, par l'intermédiaire d'une mémoire centrale, qui construit ainsi une population de solutions d'élite. La méthode proposée met en œuvre quatre méta-heuristiques : les méthodes tabou adaptées de Gendreau, Hertz et Laporte [75] et de Cordeau, Laporte et Mercier [35], ainsi que deux méthodes évolutionnaires se différenciant par l'opérateur de croisement qui est appliqué. Dans un cas il s'agit de l'opérateur OX, dans l'autre de l'opérateur ER. La population initiale est générée par des heuristiques classiques de construction (section 7.1). Chaque méthode tabou choisit une solution dans la population d'élite pour débuter la recherche et avant chaque phase de diversification, celle-ci procédant à partir de la meilleure solution entre la solution importée et la meilleure solution locale. Elle retourne à la population les nouveaux minima locaux identifiés pendant la recherche. La population des méta-heuristiques évolutionnaires est également la population d'élite. Les solutions retournées par les méta-heuristiques sont d'abord utilisées comme solutions initiales pour des méthodes de post-optimisation. Outre les heuristiques 2-opt, 3-opt et Or-opt [119], [137], les auteurs ont recours à une stratégie de type chaîne d'éjection afin d'éliminer une tournée tout en préservant l'admissibilité de la solution. Les solutions résultant sont finalement introduites dans la population. Le Bouthillier et Crainic [116] ont présenté une version de leur algoritme qui permet d'initier des phases globales d'intensification et de diversification à partir de patrons de d'arcs apparaissant dans les « bonnes » et les « mauvaises » solutions dans la memoire centrale.

L'évaluation des algorithmes est réalisée sur les 56 instances tests de Solomon [100]. Nous rappelons que tel que présenté par Solomon [172], l'objectif est tout d'abord de minimiser le nombre de tournées puis leur coût total. Si on considère donc tout d'abord le nombre de véhicules, les meilleures solutions sont celles obtenues avec les approches de Berger, Barkaoui et Bräysy [19] et Le Bouthillier et Crainic [116], ainsi que Bräysy [22] et Bent et Van Hentenryck [18], qui engendrent au total 405 tournées pour les 56 problèmes.

Les algorithmes de Homberger et Gehring [97], de Gehring et Homberger [73], de Mester [123] et de Bräysy, Hasle et Dullaert [24] génèrent 406 tournées. Cordeau, Laporte et Mercier [35] en obtiennent 407, Le Bouthillier et Crainic [115] 409, Taillard *et al.* [175] 410, Chiang et Russell [30] 411 et Rochat et Taillard [161] 415. Le coût total des solutions de Bräysy [22] est de 57272, celui de Le Bouthillier et Crainic [116] est de 57360, celui de Bent et Van Hentenryck [18] est de 57710 et celui de Berger, Barkaoui et Bräysy [19] est de 57952. Parmi les nombreuses méthodes générant un ensemble de 406 tournées, Mester [123] obtient un coût total de 57219 alors que Bräysy, Hasle et Dullaert [24] ont un coût total de 57401, Gehring et Homberger [73] de 57641 et Homberger et Gehring [97] de 57876. Cordeau, Laporte et Mercier [35] obtiennent 57556, Le Bouthillier et Crainic [115] 57574, Taillard *et al.* [175] 57523, Chiang et Russell [30] 58502 et Rochat et Taillard [161] 57231. La comparaison, au niveau temps de calcul est là encore délicate car les expériences numériques ont été conduites selon des protocoles différents sur des machines diverses. Bräysy, Hasle et Dullaert [24] donnent néanmoins une comparaison dans la mesure où les informations sont disponibles en rééchelonnant les temps par rapport à une station Sun Sparc 10. Ainsi, un temps de 486 minutes est attribué à Berger, Barkaoui et Bräysy

[19], versus 1106 minutes à Bräysy, Hasle et Dullaert [24], 1458 minutes pour Gehring et Homberger [73], 312 minutes pour Homberger et Gehring [97], 1240 minutes pour Taillard *et al.* [175] et 138 minutes pour Rochat et Taillard [161]. Le Bouthillier et Crainic [115, 116] obtiennent des accélérations linéaires par rapports aux temps sequentiels des meilleures methodes.

7.3 Le PTV en pratique

Les problèmes étudiés en recherche opérationnelle, même s'ils proviennent de cas concrets, résultent le plus souvent de simplifications de ceux-ci. Ainsi le PTV tel qu'il est abordé dans la littérature peut être considéré comme une approximation du problème auquel les transporteurs doivent faire face. Ceci est principalement dû à deux facteurs. Premièrement, les gestionnaires des entreprises de transport sont aujourd'hui très sensibles aux contraintes liées à la législation du travail. Or, les restrictions de ce type sont peu présentes dans les extensions considérées. Deuxièmement, les contraintes "classiques" telles que les contraintes de capacité ou de fenêtres de temps peuvent être pour le transporteur des contraintes "souples", c'est à dire des contraintes qui peuvent être levées en certaines occasions, alors que dans les approches de recherche opérationnelle ces contraintes doivent être vérifiées pour que la solution soit admissible. Prenons un exemple simple pour illustrer notre propos. La contrainte élémentaire qui intervient dans la plupart des variantes du PTV est la contrainte liée à la charge utile du véhicule. Cette contrainte, malgré l'apparente nécessité physique de la satisfaire, s'avère dans de nombreux cas réels une contrainte souple. Ainsi, elle ne sera respectée que de façon approximative si le responsable des transports juge par exemple que le poids des marchandises a été surévalué ou si le gain engendré par sa violation au niveau du critère optimisé est important quitte à encourir le risque d'être amendé [160].

Dès lors, toutes les heuristiques présentées ci-dessus ne sont pas pertinentes en vue d'être adaptée à la résolution de problèmes réels. Récemment Cordeau *et al.* [34] ont fait une analyse critique des différentes méthodes afin d'analyser lesquelles étaient les plus prometteuses selon cette perspective notamment. Ils ont identifié quatre critères d'évaluation des heuristiques : 1) la qualité des solutions produites en terme valeur de l'objectif mais également en termes de nombre de solutions proposées et de robustesse ; 2) le temps de calcul ; 3) la simplicité ; 4) l'adaptabilité. Si les deux premiers critères sont en général bien présents dans l'analyse des résultats, les deux seconds ne le sont pas toujours dans la conception des méthodes. Or, ils sont fondamentaux dans le développement d'approches pour la résolution de problèmes réels. Dès lors on ne peut que souhaiter que les heuristiques qui seront proposées à l'avenir soient conçues avec ce souci là.

8 Conclusions et perspectives

Le transport est aux cœur de la société humaine. On peut affirmer sans crainte de se tromper qu'il n'y a pas d'activité humaine qui ne demande le déplacement de personnes ou de biens. La recherche opérationnelle offre les méthodologies qui permettent d'accroître l'efficacité et la qualité, en termes tant économiques que de service, des services, entreprises et systèmes de transport. D'un autre coté, la complexité des opérations et des processus de planification et de gestion des systèmes de transport de marchandises offrent de grands défis et des opportunités de modélisation et algorithmiques qui font avancer la discipline.

Nous avons focalisé la présentation sur quelques problèmes et enjeux majeurs du transport de marchandises : localisation des installations et des points de rupture de charge, conception des réseaux de service et gestion des flottes pour le transport á grande distance, enfin, conception des tournées de véhicules. Dans chaque cas, nous avons présenté quelques uns des grands défis ainsi que les modèles et méthodes proposés pour les relever.

Le domaine est toujours en plein essor, nourri par, d'un côté, l'évolution de l'environnement économique, technologique, politique et social de l'industrie du transport et de la logistique et, de l'autre côté, par les progrès de l'informatique et de la recherche opérationnelle.

9 Remerciements

Ce projet a été financé en partie par le Conseil de recherches en sciences naturelles et en génie du Canada, par le biais du programme de subventions de recherche *Subventions à la découverte* et par la Communauté Européenne et la région Nord Pas de Calais, par le biais de la subvention sur les Méthodologies d'Optimisation sur les Systèmes de Transport.

10 Annexe : Méta-heuristiques

La plupart des problmes présentés dans ce chapitre sont difficiles. Les méthodes de résolution approchée, les méta-heuristiques en particulier, offrent alors la seule voie qui permettent pour des problèmes de taille réaliste d'obtenir des solutions de «bonne »qualité avec un effort de calcul «raisonable ».

La plupart des méta-heuristiques ont été employées pour résoudre des formulations reliées aux problèmes de transport, et plus particulièrement les problèmes de tournées de

véhicules. Toutefois deux d'entre elles ont conduit aux adaptations les plus performantes sur les problèmes que nous considérons dans ce chapitre. Il s'agit de la recherche tabou et des algorithmes évolutionnaires.

10.1 Recherche avec tabous

La méthode de recherche tabou (RT) est une procédure heuristique d'amélioration générale qui peut être adaptée à la résolution de nombreux problèmes d'optimisation combinatoire. Les technique de RT se sont révélées particulièrement efficaces pour obtenir de très bonnes solutions de différents problèmes tels que la conception de réseaux multimodes avec capacités, le problème du voyageur de commerce (PVC) ou le PTV. Nous présentons dans cette section les principaux éléments constitutifs de la recherche tabou. Une description plus complète a notamment été proposé dans les livres de Glover et Laguna [88], Teghem et Pirlot [177] et Dréo *et al.* [63].

La recherche tabou [85] [86] [87] est une méta-heuristique dans laquelle une procédure de recherche locale est appliquée à chaque pas d'un processus itératif de recherche générale. Plus spécifiquement, considérons le problème d'optimisation suivant :

$$\min_{s \in \mathcal{S}} f(s) \tag{99}$$

Initiée à partir d'une solution $s_0 \in \mathcal{S}$ obtenue à l'aide d'une procédure heuristique ou choisie arbitrairement dans $\mathcal{S}$, la recherche tabou explore l'espace des solutions $\mathcal{S}$ en se déplaçant d'une solution à une autre. Plus précisément, à l'itération ℓ, une nouvelle solution $s_\ell \in \mathcal{S}$ est visitée suite à une modification de la solution courante $s_{\ell-1} \in \mathcal{S}$. Une telle modification est appelée un *mouvement*. Considérant tous les mouvements admissibles depuis la solution courante $s_{\ell-1}$, nous définissons le *voisinage* de $s_{\ell-1}$, $\mathcal{V}(s_{\ell-1})$ comme le sous-ensemble de $\mathcal{S}$ associé aux solutions qui peuvent être obtenues en effectuant un mouvement.

$s_\ell \in \mathcal{V}(s_{\ell-1})$ est déterminée comme étant une des solutions voisines qui permet d'améliorer le plus la valeur $f(s_{\ell-1})$ de la fonction objectif. Si cette dernière ne peut l'être, s_ℓ peut par exemple être choisie comme étant une solution voisine de $s_{\ell-1}$ qui détériore le moins la valeur de f. Ainsi, les technique de RT évitent que le processus soit bloqué dans un minimum local. Toutefois, des restrictions de mouvement doivent être introduites afin de tenter d'éviter la réalisation d'un mouvement conduisant à une région de l'espace des solutions déjà visitée. Plus précisément, les mouvements inverses des θ derniers mouvements effectués sont interdits. Ils sont mémorisés dans une structure appelée *liste tabou* qui est mise à jour à chaque itération. θ représente la longueur de la liste. Imposer un statut tabou à un mouvement peut être trop restrictif, notamment lorsqu'un mouvement tabou amène la recherche à visiter une région a priori intéressante de S. Des conditions, appelées *conditions d'aspiration*, permettent de préciser les circonstances dans lesquelles un mouvement tabou sera accepté.

Finalement, la RT est stoppée suivant un critère d'arrêt qui peut être le nombre d'itérations réalisées, le nombre d'itérations sans amélioration de la meilleure valeur de la fonction objectif, ou encore la valeur du ratio entre la valeur de la meilleure solution et une borne inférieure sur la valeur optimale.

En résumé si nous désignons par s^*, la meilleure solution rencontrée au cours de la recherche, la version simple de la méta-heuristique tabou peut être décrite ainsi :

1. Initialiser $s^* = s_0, \ell = 0$
2. $\ell = \ell + 1$,

   Déterminer la meilleure solution $s_\ell \in V(s_{\ell-1})$ tel que le mouvement associé ne soit pas tabou ou satisfasse les conditions d'aspiration ;
3. Répéter le Pas 2 jusqu'à ce que le critère d'arrêt soit vérifié.

Plusieurs autres éléments enrichissent la méthode et en améliornet l'efficacité. La création et la mise à jour de *mémoires* emmagasinant des informations sur le déroulement de la recherche et les solutions rencontrées, et leurs attributs, constitue une caractéristique fondamentale de la méthode. Ces mémoires permettent de mettre en œuvre des restrictions tabou basés sur des critères plus complexes, de varier le choix des voisinages et d'appliquer des stratégies d'*intensification* et de *diversification* de la recherche. L'intensification consiste à explorer en détail une région a priori prometteuse de $\mathcal{S}$ où les solutions ont en commun certaines propriétés. Sa mise en oeuvre réside le plus souvent en une réduction du voisinage $V(s_{\ell-1})$ de la solution courante $s_{\ell-1}$. La diversification est une technique complémentaire à l'intensification. Son objectif est d'amener la recherche à examiner des solutions de types plus divers. Cela peut, par exemple, être réalisé en augmentant l'espace initial S des solutions ou en redémarrant la recherche à partir d'une nouvelle solution.

10.2 Algorithmes évolutionnaires

Les algorithmes évolutionnaires (AE) constituent une autre stratégie d'amélioration générale qui peut être utilisée pour résolution la nombreux problèmes d'optimisation combinatoire. Récemment, plusieurs approches basées sur les AE ont permis d'obtenir d'excellents résultats sur le PTV (voir section 7.2). Nous proposons ci-dessous une présentation générale des AE et nous renvoyons le lecteur intéressé aux livres de Davis *et al.* [58] et de Dréo *et al.* [63].

Dans les années soixante, des chercheurs dont Holland [96] ont voulu s'inspirer de mécanismes biologiques telles que les lois de Mendel et le processus naturel de sélection décrit par Darwin pour développer une méthodologie générale permettant de résoudre de façon heuristique des problèmes d'optimisation. Les AE sont donc basés sur la notion fondamentale d'évolution d'une population de solutions d'un problème.

Considérons le problème d'optimisation précédent (99). Partant d'un ensemble $\mathcal{P}_0$ de

solutions, appelée *population*, obtenues à l'aide de procédures heuristiques ou choisies arbitrairement dans S, les AE vont avoir pour but de faire évoluer cette population $\mathcal{P}_0$ vers des solutions de bonne qualité de (99). Plus précisément, à l'itération ℓ, le principe de sélection est appliqué afin de retenir qu'un sous-ensemble de solutions $\overline{\mathcal{P}}_{\ell-1}$ de $\mathcal{P}_{\ell-1}$ suivant les valeurs associées de la fonction objectif. L'idée la plus simple est d'associer à chaque solution à partir des valeurs de la fonction objectif une probabilité de sélection et de ne retenir qu'un nombre fixé de solutions en les choisissant aléatoirement suivant cette distribution de probabilité. Partant de $\overline{\mathcal{P}}_{\ell-1}$, la nouvelle population $\mathcal{P}_\ell$ est formée en ajoutant à $\overline{\mathcal{P}}_{\ell-1}$ des solutions obtenues en appliquant des opérateurs de *croissement* correspondant à l'échange de matériel génétique et de *mutation* qui représente la source de la diversité génétique. Notons que la nouvelle population $\mathcal{P}_\ell$ peut être constituée différemment. On peut par exemple inclure dans $\mathcal{P}_\ell$ uniquement des solutions résultants de l'application des opérateurs ou encore ne retenir que les meilleures solutions parmi elles.

Les algorithmes génétiques constituent une des classes des AE. Ils reposent sur le fait que les solutions sont préalablement codées sous une forme adéquate pour pouvoir appliquer les opérateurs génétiques. En fait, la similitude entre la structure génétique d'un chromosome et la représentation d'une structure complexe à l'aide d'un vecteur binaire, fut une des motivations qui poussa les chercheurs à s'inspirer de la biologie. L'opérateur de croisement vise à partir des représentations codées de deux solutions à générer une nouvelle solution en ne retenant qu'une partie du code de chacune des solutions. La solution résultante pouvant être non admissible, un opérateur de réparation peut ensuite être appliqué. Si les solutions sont codées sous la forme de vecteurs binaires, un opérateur de croissement classique, appelé *opérateur de croisement un point*, consiste à choisir aléatoire un indice i de composantes du vecteur et puis à générer une nouvelle solution en recopiant les i premiers composantes de la première solution puis les suivants de la seconde solution. L'opérateur de mutation vise à modifier une partie du codage d'une solution pour en générer une nouvelle. Comme précédemment, il peut être nécessaire d'appliquer un opérateur de réparation. Dans le cas de solutions codées comme des vecteurs binaires, un opérateur de mutation classique consiste à choisir une des composantes et à fixer sa valeur à la valeur complémentaire.

En résumé, un algorithme évolutionnaire de base peut être décrit ainsi :

1. Créer une population initiale $\mathcal{P}_0$, $\ell = 0$
2. $\ell = \ell + 1$,

   Sélectionner un ensemble de solutions $\overline{\mathcal{P}}_{\ell-1}$ de $\mathcal{P}_{\ell-1}$;

   Appliquer les opérateurs de croisement et de mutation pour former une nouvelle population $\mathcal{P}_\ell$;
3. Répéter le Pas 2 jusqu'à ce que le critère d'arrêt soit vérifié.

Plusieurs stratégies ont été proposées pour améliorer l'efficacité des AE. Les plus courantes sont l'utilisation de méthodes de recherche locale, permettant d'améliorer la population des solutions obtenues par croisement ou par mutation, et l'hybridation avec

d'autres méta-heuristiques et en particulier la recherche tabou. Dans ce dernier cas, on parle alors d'algorithmes mémétiques.

Références

[1] Achuthan, N.R., L. Caccetta, and Hill, S.P. An improved branch-and-cut algorithm for the capacitated vehicle routing problem. *Transportation Science*, 37 :153–169, 2003.

[2] Addinnour-Helm, S. and Venkataramanan M.A. Solution Approaches to Hub Location Problems. *Annals of Operations Research*, 78 :31–50, 1998.

[3] Ahuja, R.K., Magnanti, T.L., and Orlin, J.B. *Network Flows – Theory, Algorithms, and Applications*. Prentice-Hall, Englewood Cliffs, NJ., 1993.

[4] Aikens, C.H. Facility Location Models for Distribution Planning. *European Journal of Operational Research*, 22 :263–279, 1985.

[5] Armacost, A.P., Barnhart, C., and Ware, K.A. Composite Variable Formulations for Express Shipment Service Network Design. *Transportation Science*, 36(1) :1–20, 2002.

[6] Assad, A.A. Models for Rail Transportation. *Transportation Research A : Policy and Practice*, 14A :205–220, 1980.

[7] P. Augerat, Belenger, J.M., E. Benavent, A. Corberan, and Naddef D. Separating capacity constraints in the cvrp using tabu search. *European Journal of Operational Research*, 106 :546–557, 1998.

[8] Augerat Ph. *Approche polyédrale du problème de tournées de véhicules*. PhD thesis, Institut National Polytechnique de Grenoble, Grenoble, France, 1995.

[9] Aykin, T. On a Quadratic Integer Program for the Location of Interacting Hub Facilities. *European Journal of Operational Research*, 46 :409–411, 1990.

[10] Aykin, T. Lagrangian Relaxation Based Approaches to Capacitated Hub-and-Spoke Network Design Problem. *European Journal of Operational Research*, 79 :501–523, 1994.

[11] Aykin, T. Networking Policies for the Hub-and-Spoke Systems with Application to the Air Transportation System. *Transportation Science*, 29 :201–221, 1995.

[12] Aykin, T. The Hub Location and Routing Problem. *European Journal of Operational Research*, 83 :200–219, 1995.

[13] Baker, B.M. and Ayechew, M.A. A genetic algorithm for the vehicle routing problem. *Computers & Operations Research*, 30 :787–800, 2003.

[14] Balakrishnan, A., Magnanti, T.L., and Mirchandani, P. Network Design. In M. Dell'Amico, F. Maffioli, and S. Martello, editors, *Annotated Bibliographies in Combinatorial Optimization*, pages 311–334. John Wiley & Sons, New York, NY, 1997.

[15] Balinski M. and Quandt R. On an integer program for a delivery problem. *Operations Research*, 12 :300–304, 1964.

[16] Barnhart, C. and Schneur, R.R. Network Design for Express Freight Service. *Operations Research*, 44(6) :852–863, 1996.

[17] Beasley J.E. Route-first cluster-second methods for vehicle routing. *Omega*, 11 :403–408, 1983.

[18] Bent, R. and Van Hentenryck, P. A Two-Stage Hybrid Local Search for the Vehicle Routing Problem with Time Windows. *Transportation Science*, 2001. to appear.

[19] J. Berger, M. Barkaoui, and Bräysy, O. A route-directed hybrid genetic approach for the vehicle routing problem with time windows. *INFOR*, 41 :179–194, 2003.

[20] Bourbeau, B., Gendron, B., and Crainic, T.G. Branch-and-Bound Parallelization Strategies Applied to a Depot Location and Container Fleet Management Problem. *Parallel Computing*, 26(1) :27–46, 2000.

[21] Braklow, J.W., Graham, W.W., Hassler, S.M., Peck, K.E., and Powell, W.B. Interactive Optimization Improves Service and Performance for Yellow Freight System. *Interfaces*, 22(1) :147–172, 1992.

[22] Bräysy, O. A reactive variable neighborhood search algorithm for the vehicle routing problem with time windows. *INFORMS Journal on Computing*, 2001. to appear.

[23] Bräysy O. and Gendreau M. Vehicle routing problem with time windows, part ii : Metaheuristics. *Transportation Science*, à paraître.

[24] Bräysy, O., Hasle, G., and Dullaert, W. A multi-start local search algorithm for the vehicle routing problem with time windows. *European Journal of Operational Research*, 159 :586–605, 2004.

[25] Buedenbender, K., Grünert, T., and Sebastian, H.-J. A Hybrid Tabu Search/Branch and Bound Algorithm for the Direct Flight Network Design Problem. *Transportation Science*, 34(4) :364–380, 2000.

[26] Campbell, J.F. Integer Programming Formulations of Discrete Hub Location Problem. *European Journal of Operational Research*, 72 :387–405, 1994a.

[27] Campbell, J.F. Hub Location and the p-Hub Median Problem. *Operations Research*, 44(6) :923–935, 1996.

[28] Campo Dall'Orto, L., Crainic, T.G., Léal, J.E., and Powell, W.B. The Single-node Dynamic Service Scheduling and Dispatching Problem. *European Journal of Operational Research*, 2004. to appear.

[29] Cheung, R.K. and Chen, C.-Y. A Two-Stage Stochastic Network Model and Solution Methods for the Dynamic Empty Container Allocation Problem. *Transportation Science*, 32(2) :142–162, 1998.

[30] Chiang W.-C. and Russell R.A. A reactive tabu search metaheuristic for the vehicle routing problem with time windows. *INFORMS Journal on Computing*, 9 :417–430, 1997.

[31] Christiansen, M., Fagerholt, K., Nygreen, B., and Ronen, D. Maritime Transportation. In Barnhart, C. and Laporte, G., editors, *Transportation*, Handbooks in Operations Research and Management Science. North-Holland, Amsterdam, 2004.

[32] Christofides, N., Mingozzi A., and Toth, P. The Vehicle Routing Problem. In N. Christofides, Mingozzi A., P. Toth, and C. Sandi, editors, *Combinatorial Optimization*, pages 315–338. John Wiley, New York, 1979.

[33] Clarke, G. and Wright, J.W. Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research*, 12 :568–581, 1964.

[34] Cordeau, J.-F., Gendreau, M., Laporte, G., Potvin, J.-Y., and Semet, F. A Guide to Vehicle Routing Heuristics. *Journal of the Operational Research Society*, 53 :512–522, 2002.

[35] Cordeau, J.-F., Laporte, G., and Mercier, A. A Unified Tabu Search Heuristic for Vehicle Routing Problems with Time Windows. *Journal of the Operational Research Society*, 52 :928–936, 2001.

[36] Cordeau, J.-F., Toth, P., and Vigo, D. A Survey of Optimization Models for Train Routing and Scheduling. *Transportation Science*, 32(4) :380–404, 1998.

[37] R. Cordone and R. Wolfler-Calvo. A heuristic for vehicle routing problem with time windows. *Journal of Heuristics*, 7 :107–129, 2001.

[38] Crainic, T.G. Rail Tactical Planning : Issues, Models and Tools. In L. Bianco and A. La Bella, editors, *Freight Transport Planning and Logistics*, pages 463–509. Springer-Verlag, Berlin, 1988.

[39] Crainic, T.G. Network Design in Freight Transportation. *European Journal of Operational Research*, 122(2) :272–288, 2000.

[40] Crainic, T.G. Long-Haul Freight Transportation. In Hall, R.W., editor, *Handbook of Transportation Science*, pages 451–516. Kluwer Academic Publishers, Norwell, MA, second edition, 2003.

[41] Crainic, T.G. Parallel Computation, Co-operation, Tabu Search. In C. Rego and B. Alidaee, editors, *Metaheuristic Optimization Via Memory and Evolution : Tabu Search and Scatter Search*, pages 283–302. Kluwer Academic Publishers, Norwell, MA, 2005.

[42] Crainic, T.G., Dejax, P.J., and Delorme, L. Models for Multimode Multicommodity Location Problems with Interdepot Balancing Requirements. *Annals of Operations Research*, 18 :279–302, 1989.

[43] Crainic, T.G. and Delorme, L. Dual-Ascent Procedures for Multicommodity Location-Allocation Problems with Balancing Requirements. *Transportation Science*, 27(2) :90–101, 1995.

[44] Crainic, T.G., Delorme, L., and Dejax, P.J. A Branch-and-Bound Method for Multicommodity Location with Balancing Requirements. *European Journal of Operational Research*, 65(3) :368–382, 1993.

[45] Crainic, T.G., Ferland, J.-A., and Rousseau, J.-M. A Tactical Planning Model for Rail Freight Transportation. *Transportation Science*, 18(2) :165–184, 1984.

[46] Crainic, T.G., Gendreau, M., and Dejax, P.J. Dynamic Stochastic Models for the Allocation of Empty Containers. *Operations Research*, 43 :102–302, 1993.

[47] Crainic, T.G., Gendreau, M., P. Soriano, and Toulouse, M. A Tabu Search Procedure for Multicommodity Location/Allocation with Balancing Requirements. *Annals of Operations Research*, 41 :359–383, 1993.

[48] Crainic, T.G. and Kim, K.H. Intermodal Transportation. In Barnhart, C. and Laporte, G., editors, *Transportation*, Handbooks in Operations Research and Management Science. North-Holland, Amsterdam, 2005. forthcoming.

[49] Crainic, T.G. and Laporte, G. Planning Models for Freight Transportation. *European Journal of Operational Research*, 97(3) :409–438, 1997.

[50] Crainic, T.G. and Rousseau, J.-M. Multicommodity, Multimode Freight Transportation : A General Modeling and Algorithmic Framework for the Service Network Design Problem. *Transportation Research B : Methodological*, 20 :225–242, 1986.

[51] Crainic, T.G. and Roy, J. O.R. Tools for Tactical Freight Transportation Planning. *European Journal of Operational Research*, 33(3) :290–297, 1988.

[52] Crainic, T.G., Toulouse, M., and Gendreau, M. Parallel Asynchronous Tabu Search for Multicommodity Location-Allocation with Balancing Requirements. *Annals of Operations Research*, 63 :277–299, 1995.

[53] Crainic, T.G., Toulouse, M., and Gendreau, M. Synchronous Tabu Search Parallelization Strategies for Multicommodity Location-Allocation with Balancing Requirements. *OR Spektrum*, 17(2/3) :113–123, 1995.

[54] Crainic, T.G., Toulouse, M., and Gendreau, M. Towards a Taxonomy of Parallel Tabu Search Algorithms. *INFORMS Journal on Computing*, 9(1) :61–72, 1997.

[55] Dantzig, G.B. and Ramser, J.H. The truck dispatching problem. *Management Science*, 6 :80–91, 1959.

[56] Daskin, M.S. *Network and Discrete Location. Models, Algorithms, and Applications.* John Wiley & Sons, New York, NY, 1995.

[57] Daskin, M.S. and Owen, S.H. Location Models in Transportation. In Hall, R.W., editor, *Handbook of Transportation Science*, pages 321–371. Kluwer Academic Publishers, Norwell, MA, second edition, 2003.

[58] Davis, D.L., De Jong, K., Vose, M.D., and Whitley, L.D. *Evolutionary Algorithms.* Springer, New York, 1999.

[59] Dejax, P.J. and Crainic, T.G. A Review of Empty Flows and Fleet Management Models in Freight Transportation. *Transportation Science*, 21(4) :227–247, 1987.

[60] Delorme, L., Roy, J., and Rousseau, J.-M. Motor-Carrier Operation Planning Models : A State of the Art. In L. Bianco and A. La Bella, editors, *Freight Transport Planning and Logistics*, pages 510–545. Springer-Verlag, Berlin, 1988.

[61] Desrochers, M. La fabrication d'horaires de travail pour les conducteurs d'autobus par une méthode de génération de colonnes. Technical report, Publication 470, Centre de Recherche sur les Transports, Université de Montréal, Canada, 1986.

[62] Desrochers, M., J. Desrosiers, and Solomon, M.M. A new optimization algorithm for the vehicle routing problem with time windows. *Operations Research*, 40 :342–354, 1992.

[63] Dréo, J., A. Pétrowski, P. Siarry, and Taillard, É.D. *Métaheuristiques pour l'Optimisation Difficile*. Eyrolles, Paris, 2003.

[64] Z Drezner, editor. *Facility Location. A Survey of Applications and Methods*. Springer-Verlag, New York, NY, 1995.

[65] Dror, M. A mote on the complexity of the shortest path models for column generation in vrptw. *Operations Research*, 42 :977–978, 1995.

[66] L. Equi, G. Gallo, S. Marziale, and A. Weintraub. A Combined Transportation and Scheduling Problem. *European Journal of Operational Research*, 97(1) :94–104, 1997.

[67] D. Erlenkotter. A Dual–Based Procedure for Uncapacitated Facility Location. *Operations Research*, 26 :992–1009, 1978.

[68] Ernst, A.T. and Krishnamoorthy, M. Efficient Algorithms for the Uncapacitated Single Allocation p-Hub Median Problem. *Location Science*, 4(3) :139–154, 1996.

[69] Farvolden, J.M. and Powell, W.B. A Dynamic Network Model for Less-Than-Truckload Motor Carrier Operations. Working Paper 90-05, Department of Industrial Engineering, University of Toronto, Toronto, ON, Canada, 1991.

[70] Farvolden, J.M. and Powell, W.B. Subgradient Methods for the Service Network Design Problem. *Transportation Science*, 28(3) :256–272, 1994.

[71] Fisher, M.L. and Jaikumar, R. A generalized assignment heuristic for the vehicle routing. *Networks*, 11(1) :109–124, 1981.

[72] M. Franck and P. Wolfe. An Algorithm for Quadratic Programming. *Naval Research Logistics Quaterly*, 3 :95–110, 1956.

[73] Gehring, H. and Homberger, J. Parallelization of a two-phase metaheuristic for routing problems with time windows. *Journal of Heuristics*, 8 :251–276, 2002.

[74] Gendreau, M., Hertz, A., and Laporte, G. New Insertion and Postoptimization Procedures for the Traveling Salesman Problem. *Operations Research*, 40(6) :1086–1094, 1992.

[75] Gendreau, M., Hertz, A., and Laporte, G. A Tabu Search Heuristic for the Vehicle Routing Problem. *Management Science*, 40 :1276–1290, 1994.

[76] Gendreau, M., Hertz, A., Laporte, G., and Stan, M. A Generalized Insertion Heuristic for the Traveling Salesman Problem with Time Windows. *Operations Research*, 46(3) :330–335, 1998.

[77] Gendron, B. and Crainic, T.G. A Branch-and-Bound Algorithm for Depot Location and Container Fleet Management. *Location Science*, 3(1) :39–53, 1995.

[78] Gendron, B. and Crainic, T.G. A Parallel Branch-and-Bound Algorithm for Multicommodity Location with Balancing Requirements. *Computers & Operations Research*, 24(9) :829–847, 1997.

[79] Gendron, B., Potvin, J.-Y., and Soriano, P. Tabu Search with Exact Neighbor Evaluation for Multicommodity Location with Balancing Requirements. *INFOR*, 37(3) :255–270, 1999.

[80] Gendron, B., Potvin, J.-Y., and Soriano, P. A Parallel Hybrid Heuristic for the Multicommodity Capacitated Location Problem with Balancing Requirements. *Parallel Computing*, 29 :591–606, 2003.

[81] Gendron, B., Potvin, J.-Y., and Soriano, P. A Tabu Search with Slope Scaling for the Multicommodity Capacitated Location Problem with Balancing Requirements. *Annals of Operations Research*, 122 :193–217, 2003.

[82] Ghamlouche, I., Crainic, T.G., and Gendreau, M. Cycle-based Neighbourhoods for Fixed-Charge Capacitated Multicommodity Network Design. *Operations Research*, 51(4) :655–667, 2003.

[83] Ghamlouche, I., Crainic, T.G., and Gendreau, M. Path Relinking, Cycle-based Neighbourhoods and Capacitated Multicommodity Network Design. *Annals of Operations Research*, 131 :109–133, 2004.

[84] B. Gillett and L. Miller. A Heuristic Algorithm for the Vehicle Dispatch Problem. *Operations Research*, 22 :340–349, 1974.

[85] Glover, F. Future Paths for Integer Programming and Links to Artificial Intelligence. *Computers & Operations Research*, 1(3) :533–549, 1986.

[86] Glover, F. Tabu Search – Part I. *ORSA Journal on Computing*, 1(3) :190–206, 1989.

[87] Glover, F. Tabu Search – Part II. *ORSA Journal on Computing*, 2(1) :4–32, 1990.

[88] Glover, F. and Laguna, M. *Tabu Search*. Kluwer Academic Publishers, Norwell, MA, 1997.

[89] Godfrey, G.A. and Powell, W.B. An Adaptive Dynamic Programming Algorithm for Dynamic Fleet Management I : Single Period Travel Times. *Transportation Science*, 36(1) :21–39, 2002.

[90] Godfrey, G.A. and Powell, W.B. An Adaptive Dynamic Programming Algorithm for Dynamic Fleet Management II : Multiperiod Travel Times. *Transportation Science*, 36(1) :40–54, 2002.

[91] Golden, B.L., Wasil, E.A., Kelly, J.P., and Chao, I.M. Metaheuristics in Vehicle Routing. In T.G. Crainic and G. Laporte, editors, *Fleet Management and Logistics*, pages 33–56. Kluwer Academic Publishers, Boston, MA, 1998.

[92] Grünert, T. and Sebastian, H.-J. Planning Models for Long-haul Operations of Postal and Express Shipment Companies. *European Journal of Operational Research*, 122 :289–309, 2000.

[93] Grünert, T., Sebastian, H.-J., and Thäerigen, M. The Design of a Letter-Mail Transportation Network by Intelligent Techniques. In R. Sprague, editor, *Proceedings Hawaii International Conference on System Sciences 32*, 1999.

[94] E. Hadjiconstantinou, N. Christofides, and A. Mingozzi. A New Exact Algorithm for the Vehicle Routing Problem Based on q-paths and k-shortest Paths Relaxations. *Annals of Operations Research*, 61 :21–43, 1995.

[95] Haghani, A.E. Formulation and Solution of Combined Train Routing and Makeup, and Empty Car Distribution Model. *Transportation Research B : Methodological*, 23B(6) :433–452, 1989.

[96] Holland, J.H. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, 1975.

[97] Homberger, J. and Gehring, H. Two Evolutionary Metaheuristics for the Vehicle Routing Problem with Time Windows. *INFOR*, 37 :297–318, 1999.

[98] Houck, D.J., Picard, J.-C., Queyranne, M., and Vemugant, R.R. The traveling salesman as a constrained shortest path problem : Theory and computational experience. *Operations Research*, 17 :93–109, 1980.

[99] Instances de christofides, mingozzi et toth. http ://mscmga.ms.ic.ac.uk/jeb/orlib/vrpinfo.html.

[100] Instances de solomon. http ://web.cba.neu.edu/ msolomon/problems.htm.

[101] Instances de ptvc exact. http ://www.branchandcut.org.

[102] Jaillet, P., Song, G., and Yu, G. Airline Network Design and Hub Location Problems. *Location Science*, 4(3) :195–212, 1996.

[103] Jordan, W.C. and Turnquist, M.A. A Stochastic Dynamic Network Model for Railroad Car Distribution. *Transportation Science*, 17 :123–145, 1983.

[104] Jozefowiez, N. *Modélisation et résolution approchée de problèmes de tournées multi-objectifs*. PhD thesis, Université des Sciences et Technologies de Lille, France, 2004.

[105] Kim, D., Barnhart, C., Ware, K., and Reinhardt, G. Multimodal Express Package Delivery : A Service Network Design Application. *Transportation Science*, 33(4) :391–407, 1999.

[106] Klincewicz, J.G. Heuristics for the p-Hub Location Problem. *European Journal of Operational Research*, 53 :25–37, 1991.

[107] Klincewicz, J.G. Avoiding Lovcal Optima in the p-Hub Location Problem Using Tabu Search and GRASP. *Annals of Operations Research*, 40 :283–302, 1992.

[108] Klincewicz, J.G. Dual Algorithm for the Uncapacitated Hub Location Problem. *Location Science*, 4(3) :173–184, 1996.

[109] Kohl, N., Desrosiers, J., Madsen, O.B.G., Solomon, M.M., and Soumis, F. 2-path cuts for the vehicle routing problem with time windows. *Transportation Science*, 33 :101–116, 1999.

[110] Kuby, M.J. and Gray, R.G. The Hub Network Design Problem with Stopovers and Feeders : The Case of Federal Express. *Transportation Research A : Policy and Practice*, 27A(1) :1–12, 1993.

[111] Labbé, M. and Louveaux, F.V. Location Problems. In M. Dell'Amico, F. Maffioli, and S. Martello, editors, *Annotated Bibliographies in Combinatorial Optimization*, pages 261–281. John Wiley & Sons, New York, NY, 1997.

[112] Labbé, M., Peeters, D., and Thisse, J.-F. Location on Networks. In M. Ball, Magnanti, T.L., Monma, C.L., and Nemhauser, G.L., editors, *Network Routing*, volume 8 of *Handbooks in Operations Research and Management Science*, pages 551–624. North-Holland, Amsterdam, 1995.

[113] G. Laporte and Y. Nobert. Exact Algorithm for the Vehicle Routing Problem. *Annals of Discrete Mathematics*, 31 :147–184, 1987.

[114] G. Laporte, Y. Nobert, and M. Desrochers. Optimal Routing Under Capacity and Distance Restrictions. *Operations Research*, 33 :1050–1073, 1985.

[115] Le Bouthillier, A. and Crainic, T.G. A Cooperative Parallel Meta-Heuristic for the Vehicle Routing Problem with Time Windows. *Computers & Operations Research*, 32(7) :1685–1708, 2005.

[116] Le Bouthillier, A., Crainic, T.G., and Kropf, P. Towards a Guided Cooperative Search. *IEEE Intelligent Systems*, 2005. forthcoming.

[117] Leddon, C.D. and Wrathall, E. Scheduling Empty Freight Car Fleets on the Louisville and Nashville Railroad. In *Second International Symposium on the Use of Cybernetics on the Railways*, pages 1–6, 1967.

[118] Lenstra, J.K. and Rinnooy Kan, A.H.G. Complexity of vehicle routing and scheduling problems. *Networks*, 11 :221–227, 1981.

[119] Lin S. Computer solutions of the traveling salesman problem. *Bell System Technical Journal*, 44 :2245–2269, 1965.

[120] Liu F.H. and Shen S.Y. A route-neighborhood-based metaheuristic for vehicle routing problem with time windows. *European Journal of Operational Research*, 118 :485–504, 1999.

[121] Lysgaard, J., Letchford, A.N., and Eglese R.W. A new branch-and-cut algorithm for the capacitated vehicle routing problem. *Mathematical Programming*, 100 :423–445, 2004.

[122] Magnanti, T.L. and Wong, R.T. Network Design and Transportation Planning : Models and Algorithms. *Transportation Science*, 18(1) :1–55, 1984.

[123] Mester, D. An evolutionary strategies algorithm for large scale vehicle routing problem with capacitate and time windows restrictions. Technical report, Working Paper, Institute of Evolution, University of Haifa, Israel, 2002.

[124] Minoux, M. Network Synthesis and Optimum Network Design Problems : Models, Solution Methods and Applications. *Networks*, 19 :313–360, 1989.

[125] Mirchandani, P.S. and Francis, R.L., editors. *Discrete Location Theory*. John Wiley & Sons, New York, NY, 1990.

[126] S.C. Misra. Linear Programming of Empty Wagon Disposition. *Rail International*, 3 :151–158, 1972.

[127] Mole, R.H. and Jameso, S.R. A sequential route-building algorithm employing a generalized savings criterion. *Operational Research Quarterly*, 27 :503–511, 1976.

[128] D. Naddef and G. Rinaldi. The graphical relaxation : A new framework for the symmetric traveling salesman polytope. *Mathematical Programming*, 58 :53–88, 1993.

[129] D. Naddef and S. Thienel. Efficient separation routines for the symmetric traveling salesman problem i : General tools and comb separation. *Mathematical Programming*, 92 :237–255, 2002.

[130] D. Naddef and S. Thienel. Efficient separation routines for the symmetric traveling salesman problem ii : Separating multi handle inequalities. *Mathematical Programming*, 92 :257–285, 2002.

[131] Nemhauser, G.L. and Wolsey, L.A. *Integer and Combinatorial Optimization*. Wiley, New York, NY, 1988.

[132] O'Kelly, M.E. A Quadratic Integer Program for the Location of Interacting Hub Facilities. *European Journal of Operational Research*, 32 :393–404, 1987.

[133] O'Kelly, M.E. A Clustering Approach to the Planar Hub Location Problem. *Annals of Operations Research*, 40 :339–353, 1992.

[134] O'Kelly, M.E. Hub Facilities Location with Fixed Costs. *Papers in Regional Science*, 71 :292–306, 1992.

[135] O'Kelly, M.E., Bryan, D., Skorin-Kapov, D, and Skorin-Kapov, J. Hub Network Design with Single and Multiple Allocation : A Computational Study. *Location Science*, 4(3) :125–138, 1996.

[136] O'Kelly, M.E., Skorin-Kapov, D, and Skorin-Kapov, J. Lower Bounds for the Hub Location Problem. *Management Science*, 41 :713–721, 1995.

[137] Or I. *Traveling salesman-type combinatorial problems and their relation to the logistics of regional blood banking*. PhD thesis, Northwestern University, Evanston, U.S.A., 1976.

[138] Osman, I.H. Metastrategy simulated annealing and tabu search for the vehicle routing problem. *Annals of Operations Research*, 41 :421–451, 1993.

[139] H. Pirkul and V. Jayaraman. Production, Transportation, and Distribution Planning in a Multi-Commodity Tri-Echelon System. *Transportation Science*, 30(4) :291–302, 1996.

[140] H. Pirkul and V. Jayaraman. A Multi-Commodity, Multi-Plant, Capacitated Facility Location Problem : Formulation and Efficient Heuristic Solution. *Computers & Operations Research*, 25(10) :869–878, 1998.

[141] Potvin J.-Y. and Rousseau J.-M. An exchange heuristic for routing problems with time windows. *Journal of the Operational Research Society*, 46 :1433–1446, 1995.

[142] Powell, W.B. A Local Improvement Heuristic for the Design of Less-than-Truckload Motor Carrier Networks. *Transportation Science*, 20(4) :246–357, 1986.

[143] Powell, W.B. A Comparative Review of Alternative Algorithms for the Dynamic Vehicle Allocation Problem. In B.L. Golden and A.A. Assad, editor, *Vehicle Routing : Methods and Studies*, pages 249–292. North-Holland, Amsterdam, 1988.

[144] Powell, W.B. Dynamic Models of Transportation Operations. In Graves, S. and Tok, T.A.G., editors, *Supply Chain Management*, volume 11 of *Handbooks in Operations Research and Management Science*, pages 677–756. North-Holland, Amsterdam, 2003.

[145] Powell, W.B., Bouzaïene-Ayari, B., and Simaõ, H.P. Dynamic Models for Freight Transportation. In Barnhart, C. and Laporte, G., editors, *Transportation*, Handbooks in Operations Research and Management Science. North-Holland, Amsterdam, 2005. to appear.

[146] Powell, W.B. and Carvalho, T.A. Dynamic Control of Multicommodity Fleet Management Problems. *European Journal of Operations Research*, 98 :522–541, 1997.

[147] Powell, W.B. and Carvalho, T.A. Dynamic Control of Logistics Queueing Networks for Large-Scale Fleet Management. *Transportation Science*, 32(2) :90–109, 1998.

[148] Powell, W.B. and Carvalho, T.A. Real-Time Optimization of Containers and Flatcars for Intermodal Operations. *Transportation Science*, 32(2) :110–126, 1998.

[149] Powell, W.B., Carvalho, T.A., Godfrey, G.A., and Simaõ, H.P. Dynamic Fleet Management as a Logistics Queueing Network. *Annals of Operations Research*, 61 :165–188, 1995.

[150] Powell, W.B., Jaillet, P., and Odoni, A. Stochastic and Dynamic Networks and Routing. In Ball, M., Magnanti, T.L., Monma, C.L., and Nemhauser, G.L., editors, *Network Routing*, volume 8 of *Handbooks in Operations Research and Management Science*, pages 141–295. North-Holland, Amsterdam, 1995.

[151] Powell, W.B. and Sheffi, Y. The Load-Planning Problem of Motor Carriers : Problem Description and a Proposed Solution Approach. *Transportation Research A : Policy and Practice*, 17(6) :471–480, 1983.

[152] Powell, W.B. and Sheffi. Y. Interactive Optimization for Motor Carrier Load Planning. *Journal of Business Logistics*, 7(2) :64–90, 1986.

[153] Powell, W.B. and Sheffi, Y. Design and Implementation of an Interactive Optimization System for the Network Design in the Motor Carrier Industry. *Operations Research*, 37(1) :12–29, 1989.

[154] Powell, W.B. and Topaloglu, H. Stochastic Programming in Transportation and Logistics. In Ruszczynski, A. and Shapiro, A., editors, *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 555–635. North-Holland, Amsterdam, 2003.

[155] Powell, W.B. and Topaloglu, H. Fleet Management. In Wallace, S. and Ziemba, W., editors, *Applications of Stochastic Programming*, Math Programming Society - SIAM Series on Optimization. SIAM, 2005. to appear.

[156] Prins, C. A simple and effective evolutionary algorithm for the vehicle routing problem. *Computers & Operations Research*, 31 :1985–2002, 2004.

[157] Ralphs, T.K., L. Kopman, Pulleyblank, W.R., and Trotter L.E. On the capacitated vehicle routing problem. *Mathematical Programming*, 94 :343–359, 2003.

[158] Rego, C. and Roucairol, C. A Parallel Tabu Search Algorithm Using Ejection Chains for the VRP. In I.H. Osman and J.P. Kelly, editors, *Meta-Heuristics : Theory & Applications*, pages 253–295. Kluwer Academic Publishers, Norwell, MA, 1996.

[159] J. Renaud, Boctor, F.F., and G. Laporte. An improved petal heuristic for the vehicle routing problem. *Journal of the Operational Research Society*, 47 :1156–1167, 1996.

[160] Rochat, Y. and Semet, F. A tabu search approach for delivering pet food and floor in switzerland. *Journal of the Operational Research Society*, 45 :1233–1246, 1994.

[161] Rochat, Y. and Taillard, É.D. Probabilistic Diversification and Intensification in Local Search for Vehicle Routing. *Journal of Heuristics*, 1(1) :147–167, 1995.

[162] Roy, J. *Un modèle de planification globale pour le transport routier des marchandises*. PhD thesis, Ecole des Hautes Etudes Commerciales, Université de Montréal, Montréal, Canada, 1984.

[163] Roy, J. and Crainic, T.G. Improving Intercity Freight Routing with a Tactical Planning Model. *Interfaces*, 22(3) :31–44, 1992.

[164] Roy, J. and Delorme, L. NETPLAN : A Network Optimization Model for Tactical Planning in the Less-than-Truckload Motor-Carrier Industry. *INFOR*, 27(1) :22–35, 1989.

[165] Russell, R.A. An effective heuristic for the m-tour traveling salesman problem with some side conditions. *Operations Research*, 25 :517–524, 1977.

[166] Russell, R.A. Hybrid heuristics for the vehicle routing problem with time windows. *Transportation Science*, 29 :156–166, 1995.

[167] Salkin, H.M. and Mathur, K. *Foundations of Integer Programming*. North-Holland, Amsterdam, 1989.

[168] Shaw, P. Using constraint programming and local search methods to solve vehicle routing problems. In M. Maher and J.-F. Puget, editors, *Principles and Practice of Constraint Programming*, Lecture Notes in Computer Science, pages 417–431. Springer-Verlag, New York, 1998.

[169] D. Skorin-Kapov and J. Skorin-Kapov. On Tabu Search for the Location of Interacting Hub Facilities. *European Journal of Operational Research*, 73 :502–509, 1994.

[170] Skorin-Kapov, D., Skorin-Kapov, J., and O'Kelly, M.E. Tight Linear Programming relaxation of Uncapacitated p-Hub Median Problems. *European Journal of Operational Research*, 94 :582–593, 1996.

[171] Smith, K., Krishnamoorthy, M., and Palaniswami, M. Neural Versus Traditional Approaches to the Location of Interacting Hub Facilities. *Location Science*, 4(3) :155–171, 1996.

[172] Solomon, M.M. Time Window Constrained Routing and Scheduling Problems. *Operations Research*, 35 :254–265, 1987.

[173] Solomon, M.M., E. Baker, and J. Schaffer. *Vehicle Routing : Methods and Studies*, chapter Vehicle Routing and Scheduling with Time Windows : Efficient Implementations of Solution Improvement Procedures, pages 85–106. North-Holland, 1988.

[174] Taillard, É.D. Parallel Iterative Search Methods for Vehicle Routing Problems. *Networks*, 23 :661–673, 1993.

[175] Taillard, É.D., Badeau, P., Gendreau, M., Guertin, F., and Potvin, J.-Y. A Tabu Search Heuristic for the Vehicle Routing Problem with Soft Time Windows. *Transportation Science*, 31(2) :170–186, 1997.

[176] Tarantilis, C.D. and Kiranoudis, C.T. Bone route : An adaptive memory-based method for effective fleet management. *Annals of Operations Research*, 115 :227–241, 2002.

[177] Teghem, J. and Pirlot, M. *Résolution de Problèmes de RO par les Métaheuristiques.* Hermès, Paris, 2003.

[178] P. Thompson and H. Psaraftis. Cyclic transfer algorithms for multivehicle routing and scheduling problems. *Operations Research*, 41 :935–946, 1993.

[179] P. Toth and D. Vigo, editors. *The Vehicle Routing Problem*, volume 9 of *SIAM Monographs on Discrete Mathematics and Applications*. SIAM, 2002.

[180] P. Toth and D. Vigo. The granular tabu search and its application to the vehicle routing problem. *INFORMS Journal on Computing*, 15 :333–348, 2003.

[181] Van Breedam, A. Improvement heuristics for the vehicle routing problem based on simulated annealing. *European Journal of Operational Research*, 86 :480–490, 1995.

[182] W.W. White. Dynamic Transshipment Networks : An Algorithm and its Application to the Distribution of Empty Containers. *Networks*, 2(3) :211–236, 1972.

[183] White, W.W. A Program for Empty Freight Car Allocation. Technical Report 360D.29.002, IBM Contributed Program Library, IBM Corporation, Program Information Department, Hawthorne, N.Y., 1968.

[184] White, W.W. and Bomberault, A.M. A Network Algorithm for Empty Freight Car Allocation. *IBM Systems Journal*, 8(2) :147–171, 1969.